



Název: Výzkum diskrétních voleb  
Podnázev: Průvodce designem experimentu, specifikací modelu a jeho odhadem se skripty pro R  
Autor: Michal Šimeček  
Vydavatel: Centrum dopravního výzkumu, v. v. i.  
Recenzovali: Paul Riley (JASPERS), Stanislav Ježek (FSS MU)  
Rok vydání: 2019  
ISBN: 978-80-88074-64-9

Tato kniha vznikla za finanční podpory Ministerstva školství, mládeže a tělovýchovy v rámci programu Národní program udržitelnosti I, projektu Dopravní VaV centrum (LO1610) na výzkumné infrastruktuře pořízené z Operačního programu Výzkum a vývoj pro inovace (CZ.1.05/2.1.00/03.0064).

ISBN: 978-80-88074-64-9



# Obsah

1.	Úvod.....	7
1.1	Historie výzkumu rozhodování .....	7
1.2	Měření rozhodování .....	9
1.3	Volba z hlediska analýzy a modelování diskretních voleb .....	10
1.4	Data pro analýzu diskretních voleb .....	11
2.	Modely rozhodování.....	13
2.1	Základní model rozhodování .....	13
2.2	Princip nezávislosti irelevantních alternativ .....	17
2.3	Modely s vnořenou strukturou – nested logit.....	18
2.4	Model pro agregované alternativy.....	19
2.5	Smíšené modely.....	22
2.6	Software pro analýzu diskretních voleb.....	23
3.	Tvorba experimentu.....	24
3.1	Příprava scénáře .....	24
3.2	Design experimentu.....	25
3.3	Personalizace experimentu.....	28
3.4	Adaptivní testování.....	28
3.5	Výběr respondentů .....	31
3.6	Pilotní studie.....	32
4.	Analýza diskretních voleb .....	33
4.1	Odhad parametrů a hodnocení kvality modelu.....	33
4.2	Kvantifikace parametrů .....	35
4.3	Elasticita .....	41
4.4	Interpretace koeficientů u smíšených modelů .....	45
5.	Případová studie: Výběr pronájmu studentského bydlení.....	46
5.1	Příprava experimentu .....	46
5.2	Výběrový soubor.....	47
5.3	Příprava dat.....	47
5.4	Základní model .....	48
5.5	Smíšený model.....	50
5.6	Kvantifikace parametrů .....	52
5.7	Shrnutí.....	54

6.	Případová studie: volba dopravního prostředku na cestu po brněnské metropolitní oblasti (BMO) .....	55
6.1	Příprava experimentu .....	56
6.2	Výběrový soubor.....	57
6.3	Příprava dat.....	58
6.4	Model pro celková data.....	60
6.5	Kvantifikace parametrů modelu pro celková data .....	64
6.6	Modely pro Brňany a obyvatele zázemí Brna .....	65
6.7	Kvantifikace parametrů modelů pro Brňany a obyvatele zázemí Brna .....	67
6.8	Elasticita volby dopravního módu - inkrementální model.....	68
6.9	Shrnutí.....	71
7.	Závěr .....	73
7.1	Kombinace <i>RP</i> a <i>SP</i> dat .....	73
7.2	Modely diskrétních voleb v dopravních modelech .....	74
7.3	Další studium .....	75
	Literatura .....	77

## Předmluva

Téměř všechno, co lidé dělají, můžeme chápat jako proces volby. Lidé si neustále vybírají mezi různými alternativami a toto chování se již před časem stalo předmětem zkoumání jak ekonomie, tak i dalších společenských věd. Byla vytvořena řada nástrojů, jak rozhodování měřit a jak jej analyzovat. Tato publikace se zabývá úvodem do designu experimentální metody vyjádřených preferencí (SP – stated preference) a do analýzy diskretních voleb (DCA – discrete choice analysis). SP experimenty jsou jedním ze způsobů, jak aspekty rozhodování měřit a analýza diskretních voleb je sada statistických modelů a technik, pomocí kterých lze výsledky měření zpracovat a interpretovat je.

Výzkum lidského rozhodování má řadu aplikací od marketingu přes dopravu, zdravotnictví až po ekologickou politiku. Tato publikace představuje úvod do aplikovaného výzkumu rozhodování obecně, jen s o něco větším důrazem na jeho využití v dopravě. Krátce se budeme věnovat historii výzkumu rozhodování a nastíníme si některé metodologické předpoklady a otázky (kapitola 1). Představíme si základní model rozhodování (kapitola 2) a poté se budeme v knize zabývat jedním z široce využívaných nástrojů výzkumu volby, kterým je zmíněný experiment (nebo též jen průzkum) vyjádřených preferencí. Ukážeme si, jak takový experiment připravit a jak jej realizovat (kapitola 3). Následně si představíme základní možnosti analýzy získaných dat pomocí analýzy a modelování diskretních voleb a interpretace výsledků (kapitola 4), kde jsou diskutovány zejména aplikace v dopravě. Na dvou kazuistických případech použití metody vyjádřených preferencí a analýz diskretních voleb si ukážeme tvorbu designu experimentu a analýzu výsledků včetně interpretace. V případové studii volby studentského bydlení (kapitola 5) si ukážeme zpracování panelových dat a interpretaci výsledků analýz pomocí kvantifikace odhadnutých parametrů na měsíčně vynaloženou finanční hodnotu. V případové studii volby způsobu dopravy respondentů z brněnské metropolitní oblasti (kapitola 6) si ukážeme použití modelu s vnořenou strukturou, kvantifikaci parametrů modelu s výpočtem hodnoty času a modelování změny využití ve způsobu dopravy při změně ceny veřejné dopravy pomocí inkrementálního modelu. V závěrečné kapitole probereme možnost kombinování různých druhů dat v jednom modelu a krátce zmíníme, jak se výsledky analýz rozhodování projevuje v dopravních modelech.

Knihou předpokládá čtenáře, který je zorientovaný v základech vědeckého výzkumu, ovládá matematiku na středoškolské úrovni a má povědomí o statistice. Knihou obsahuje na několika místech matematické vzorce. V některých případech není nezbytné, aby jim čtenář plně porozuměl, a proto se jich není potřeba děsit. Některé kapitoly téma rozšiřují nad rámec základů analýzy diskretních voleb. Například porozumění agregovaným alternativám je nezbytné pro modelování volby destinace cesty v dopravním modelování, ale v mnoha jiných aplikacích se s tímto problémem čtenář vůbec nemusí setkat.

Knihou se odvolává na datasey a skripty napsané pro statistický nástroj R (více o tomto i dalších softwarech naleznete v kapitole 2.6). Důležité části skriptů v R jsou také přímo v textu. Tyto

skripty představují praktický úvod do analýzy diskretních voleb a po drobných modifikacích poslouží též k analýze jiných dat, než jaká jsou probírána v této publikaci. Autor knihy má původně zkušenost s programem Biogeme a skripty v R připravil jen pro tuto příležitost. Nyní však začal R více používat ke své práci a tyto skripty mu ji usnadňují.

Po přečtení knihy bude čtenář schopen sám připravit jednoduchý výzkum vyjádřených preferencí, bude jej umět analyzovat pomocí základních modelů diskretních voleb a interpretovat základní výsledky. Pokud se pak čtenáři dostanou do ruky zprávy o jiných podobných výzkumech, bude jim schopen porozumět a posoudit jejich užitečnost. Bez dalšího studia se čtenář patrně nestane expertem na výzkumy rozhodování, ale stane se informovaným příjemcem výsledků těchto výzkumů. V závěrečné kapitole (kapitola 7) najde čtenář odkazy na některou rozšiřující literaturu, která mu umožní prohlubovat dál své znalosti, a seznámí se také s pohledem na dopravní modelování jako na sadu modelů diskretních voleb.

# 1. Úvod

V úvodní kapitole se budeme krátce zabývat historií výzkumu rozhodování a probereme některé metodologické otázky spojené s tímto výzkumem. Podíváme se také, jak je na lidské rozhodování nahlíženo v různých vědních oborech a jakými různými způsoby se měří jeho charakteristiky.

## 1.1 Historie výzkumu rozhodování

Výzkum rozhodování a modelování rozhodování vzniklo nezávisle v oblasti matematické psychologie a ekonometrie. Ve 20. letech dvacátého století formuloval psychometrik Louis Leon Thurstone pravidlo srovnávacího úsudku, což byla teorie aplikovatelná nejen na rozlišení dvojice psychofyzických podnětů, ale také přímo na rozhodování (Thurstone, 1994). Thurstonův model pracoval s neurčitostí v rozlišování mezi dvěma podněty, což později v teorii rozhodování našlo uplatnění v podobě teorie náhodného užítku Daniela McFaddena (McFadden, 1986).

První matematické modely rozhodování byly součástí ekonomické teorie, a zabývaly se rozhodováním jednotlivých ekonomických aktérů. Tyto modely vycházely z představ neoklasické ekonomie, podle které se jedinci nebo organizace rozhodují tak, aby maximalizovali svůj užitek. Hovoří se o teorii maximalizace užítku nebo také o teorii racionální volby. Rozhodování má být podle této teorie racionální, přičemž aktéři mají mít k dispozici všechny relevantní informace. Za těchto podmínek je procesu rozhodování možné porozumět, modelovat jej a předpovídat budoucí výsledky. Každé z alternativ, které přicházejí do úvahy lze na základě jejich vlastností přiřadit užitek, který by aktér měl z přijetí té které alternativy. Při maximalizaci užítku pak aktér volí takovou z alternativ, která nabízí užitek největší.

Jak se hledaly aplikace teorie rozhodování, ukazovalo se stále více, že původní požadavky na teorii nebudou udržitelné. John von Neumann a Oskar Morgenstern (2007) přichází v roce 1947 s aplikací teorie maximalizace užítku v teorii her, tedy v situaci, kdy nejsou k dispozici všechny potřebné informace o alternativách volby. Tyto informace jsou totiž závislé na dosud neznámých rozhodnutích ostatních hráčů (von Neumann & Morgenstern, 2007). Aktér může pouze odhadovat očekávaný užitek, který je násobkem pravděpodobnosti určitého výsledku a užítku z něj. Díky práci von Neumanna našly modely volby své uplatnění v optimalizačních procesech a strojovém učení.

Z hlediska lidské teorie rozhodování se ukázala teorie racionální volby jako příliš normativní, a to i ve verzi teorie očekávaného užítku. To znamená, že teorie racionální volby představuje jakýsi ideál, podle kterého se ovšem lidé nerozhodují. Kahneman a Tversky (1979) přichází s teorií prospektu, která lidské rozhodování popisuje a ukazuje rozdíly mezi ním a teorií očekávaného užítku. Reálně člověk při rozhodování používá heuristiky, které mu rozhodování usnadňují. Na principech teorie prospektu lze postavit různé marketingové „chytáky“,



kteře způsobí, že se lidé nemohou chovat z ekonomického hlediska racionálně. Proto jsou principy teorie prospektu někdy vnímány jako projevy lidského chybování. Gigerenzer naopak zdůrazňuje, že heuristiky usnadňují rozhodování a v obvyklých případech vedou k dostatečně dobrým výsledkům (Gigerenzer & Goldstein, 1996), hovoří o ekologické validitě těchto heuristik. Lidská mysl není podle Gigerenzera racionální ve smyslu teorie racionální volby a ani nepodléhá zkreslením, jak by se dalo vyčíst z prací Kahnemana a Tverského, ale její racionalita je omezená časem a výpočetním výkonem, který má člověk k dispozici. Cílem výzkumu rozhodování pak není zkoumání heuristik, pouze je potřeba počítat s jejich přítomností. V praxi se ukazuje, že heuristiky mohou snadno ovlivnit výsledky experimentů, přitom je obtížné rozhodnout, zdali jsou součástí zkoumaného rozhodovacího procesu nebo artefaktem, způsobeným experimentální situací.

Daniel McFadden v 70. letech přichází s teorií náhodného užitku. Podle této teorie nemůžeme jednoznačně modelovat skutečná rozhodování lidí, protože vždy může existovat nějaká komponenta rozhodování, kterou neznáme. Náhodný užitek je právě touto neznámou. McFadden pracoval na statistické analýze rozhodování a vytvořil řadu matematických modelů a dalších statistických nástrojů, pomocí kterých je možné rozhodování měřit a analyzovat. Výsledky průzkumů a analýz rozhodování bylo tedy možné použít pro zjištění preferencí zákazníků různých služeb a nastavit tyto služby tak, aby zákazníkům více vyhovovaly (Dubin & McFadden, 1984; Train et al., 1987). Moshe Ben-Akiva pomohl výzkum rozhodování rozšířit zejména do dopravního výzkumu, a to v oblasti dopravní poptávky. Výsledky výzkumu rozhodování v poptávce po dopravě se uplatňují v dopravním modelování, kde se používají k modelování volby cíle cesty, dopravního prostředku, i volbě trasy (Ben-Akiva & Lerman, 1985; Ben-Akiva & Bierlaire, 1999). Zároveň je tato analýza dopravní poptávky důležitým zdrojem dat pro plány udržitelné mobility (Loo, 2002) a kvantifikaci užitku z ekonomicky posuzovaných staveb dopravní infrastruktury (JASPERS, 2014).

Výzkum rozhodování se stal součástí celé řady oblastí aplikovaných věd. Uvedme jen pár příkladů, kde se rozhodování zkoumá a na jaké otázky může výzkum dávat odpovědi:

- Marketing – Co rozhoduje při výběru určitého produktu (služby)?
- Politika – Jaká je ochota platit za určitou veřejnou službu?  
– Jaké vlastnosti má mít kandidát, kterému voliči dají ve volbách přednost?
- Sociální geografie – Co rozhoduje o migraci obyvatelstva mezi oblastmi?
- Zaměstnanost – Co rozhoduje při výběru zaměstnání?
- Zdravotnictví – Jaké způsoby léčby pacienti preferují?
- Investování – Co rozhoduje o investicích (délka, výnosy, rizika...)?
- Doprava – Jaká je hodnota času v dopravě?  
– Jaký podíl v dopravě by získalo vysokorychlostní železniční spojení Praha – Brno?

- Jak se změní podíl využívání železnice při zvýšení rychlosti po rekonstrukci tratě?
- Enviromentalistika – jaký přístup ke krajině je obyvatelstvo ochotné financovat ze svých daní?
- Sociální práce – jak nastavit sociální službu tak, aby byla pro klienty co nejvíc akceptovatelná?
- Jaké signály vedou svědky katastrofy k jakému protektivnímu jednání?







Podrobnější informace o historii výzkumu rozhodování lze nalézt v dalších publikacích (Mansky, 1981; McFadden, 1982; Ben-Akiva & Lerman, 1985). Možnosti uplatnění výzkumu rozhodování jsou nepřehledné a na internetu lze najít řadu studií z nejrůznějších oblastí lidského života.

## 1.2 Měření rozhodování

Existuje více zdrojů dat, které lze pro výzkum rozhodování použít. Mohou to být údaje o skutečně proběhlých rozhodnutích nebo data vzniklá dotazováním některou z dotazovacích nebo experimentálních metod.

Můžeme se ptát respondentů na motivace určitých voleb. Můžeme se například ptát: „Co vás vedlo ke koupi výrobku X? Proč jste mu dal/a přednost před Y?“ Takové dotazování má několik potíží. Rozhodovací proces samotný může být pro respondenta nevědomý (Martin & Morich, 2011) a to, co se od něj dozvíme, bude jeho dodatečná interpretace vlastního chování.

Jinou možností je předložit respondentům určité alternativy a nechat je posoudit jejich atraktivitu. Čtenář například mohl slyšet o *conjoint* analýze (Green & Rao, 1971), která se uplatňuje především v marketingu. Conjoint analýza používá právě takový způsob odpovědi, tedy zhodnotit na lickertovské škále atraktivitu každé alternativy. Popřípadě nechává respondenta alternativy podle atraktivity alespoň seřadit.

Operační systém			
Rozlišení displeje	480x800	600x1024	768x1366
Rozlišení fotoaparátu	4 MPx	8 MPx	6 MPx
Velikost			
Cena	6 000 Kč	4 000 Kč	2 000 Kč

Obrázek 1: Ukázka scénáře pro conjoint analýzu

Conjoint analýza se musí vyrovnávat s několika obtížemi. Aby bylo možné postihnout vztahy mezi vlastnostmi produktů (v našem případě např. velikost nebo cena), musí respondent

pracovat s velkým počtem produktů najednou nebo postupně. To je kognitivně náročné a respondenti se budou snažit si řešení nějak usnadnit, což povede ke zkreslením. Úloha, která spočívá v hodnocení užitečnosti alternativ nebo k jejich řazení vede u stejných respondentů k jiným výsledkům, než volba nejvíce preferované alternativy (Merino-Castello, 2003). Také se může projevit efekt úlohy, který jsme poznali v jiných oblastech experimentální psychologie. Odpovědi respondentů jsou do určité míry závislé na konkrétní úloze, kterou dostanou. Kvalita odpovědí záleží na tom, jestli je zadání úlohy pro respondenty obvyklé nebo nikoli. Pokud se jedná o typ úlohy, se kterým má respondent dřívější zkušenosti a je pro něj obvyklá v jeho každodenním životě, budou výsledky takové úlohy relativně přesné a konzistentní (Šikl & Šimeček, 2015). Úloha porovnávat výrobky se více blíží každodenní zkušenosti respondenta, než když je úkolem pro každý výrobek vyjadřovat míru preference. Nejvíce se však respondentově zkušenosti podobá experiment, ve kterém má podle vlastních preferencí respondent zvolit jeden nejlepší produkt z několika. Proto vznikla varianta *conjoint* designu, založená na respondentově volbě. Tato metoda (*choice based conjoint*) je již jiným označením pro SP experiment, nebo také *DCE – Discrete Choice Experiment* (Carson & Louviere, 2011). Je to standardní a v současné době užívaná metoda výzkumu volby mezi hypotetickými alternativami, službami nebo produkty (Ewing & Sarigöllü, 2000; Bunch et al., 1993). Výsledky z SP experimentu a dalších metod zkoumajících rozhodování se zpracovávají pomocí analýzy diskrétních voleb (*DCA – Discrete Choice Analysis*). V rámci DCA je původce rozhodnutí (*decision maker*) modelován jako ten, kdo z množiny alternativ, které jsou pro výběr k dispozici, vybírá alternativu s největším užitekem (Ben-Akiva & Lerman, 1985). Vytváření SP experimentu a analýzy výsledků se budeme věnovat dále v této knize.




### 1.3 Volba z hlediska analýzy a modelování diskrétních voleb

Analýza diskrétních voleb předpokládá, že rozhodování je proces, v jehož závěru je vybrána (realizována) jedna z konečného počtu relevantních alternativ, vzájemně oddělených (exkluzivních neboli diskrétních), které jsou popsány pozorovatelnými vlastnostmi (nezávislými proměnnými, které se nazývají atributy). Principiálně musí být možné, aby byla vybrána kterákoli z alternativ, nebo přesněji, aby pravděpodobnost jejího výběru nebo nevyběru závisela také na jejích attributech. Zjistit vliv těchto atributů pro rozhodování je pak hlavním cílem analýzy.

Mnozí autoři (Aarts, Verplanken & van Knippenberg, 1998; Henscher, Rose, & Greene, 2007) nemají problém s tím, že zdánlivé rozhodnutí může být ve skutečnosti projevem zvyku, nápodoby (Chartrand, 2005; Tanner et al., 2007) nebo dalších procesů, které se rozhodováním v pravém slova smyslu být nezdají. Obecně lze říci, že každý proces, v rámci kterého se realizuje právě jedna z konečného počtu diskrétních možností, může být považován za situaci diskrétní volby. Podívejme se na tento poměrně extrémní případ. Statistiky o obyvatelstvu udávají, kolik obyvatel se v určitém roce přestěhovalo z jedné do druhé obce na území ČR. Vorel a Maier (2007) použili k modelování této vnitrostátní migrace obyvatelstva model diskrétní volby. Jako parametry tohoto modelu uvažovali sociálně-

geografické vlastnosti území. Nikdo nepochybuje o tom, že každý z obyvatel se stěhoval z nějakých svých a dost možná osobních důvodů, které mají s obecnými charakteristikami území jen málo společného. Ať už ale motivace a konkrétní preference každého z obyvatel (ti, kteří se nikam nestěhovali, také udělali volbu – rozhodli se nestěhovat) byly jakékoli, celkově nějak souvisely s obecnými nezávislými proměnnými, které Vorel a Maier sledovali a použili je při formulaci „rozhodovacího“ modelu migrace. Například je při rozhodování o bydlišti sousedství otevřených prostor preferováno více než sousedství městského parku. Nejvíce preferovaná časová dostupnost obchodu se základním zbožím je do 3 minut, zatímco časová dostupnost zastávky veřejné dopravy je nejvíce preferovaná do 7 minut od bydliště.

Každý proces volby je rámovaný scénářem. Ve scénáři je popsán konečný počet vzájemně exkluzivních alternativ, mezi kterými se jedna vybírá. Scénář může obsahovat nejen předkládané alternativy, ale i další podmínky.

Predstavte si, že cestujete do zamestnaní lebo na služobnú cestu 20 km ďaleko po územíu Slovenska zo svojho bydliska. Tu máte rôzne možnosti, ako tu cestu podniknúť.		
	<b>VLAK</b>	
Doprava z domu na vlak	15 min	
Doba jazdy vlakom	0h 20min	
Doprava z vlaku	15 min	
<b>Celková doba prepravy</b>	<b>0h 50min</b>	
Cena základného jazdného	1.60 €	
Vy uplatníte zľavu, ak máte		
Obsadenosť	istota sedení	
	Moderný voz s klimatizáciou a wi-fi	
	<b>AUTOBUS</b>	
Doprava z domu na autobus	10 min	
Doba jazdy autobusom	0h 15min	
Doprava z autobusa	15 min	
<b>Celková doba prepravy</b>	<b>0h 40min</b>	
Cena základného jazdného	1.30 €	
Vy uplatníte zľavu, ak máte		
Obsadenosť	riziko stání	
	Starý voz	
	<b>AUTOMOBIL</b>	
Doba jazdy	0h 20min	
Cena za pohonné hmoty a parkovanie	1.60 €	

Obrázek 2: Ukázka scénáře SP experimentu průzkumu v rámci studie realizovatelnosti modernizace železničního koridoru Žilina – Košice – Čierna n. T. (Star EU & CDV, 2014).

Na obrázku 2 je scénář SP experimentu, zaměřeného na výběr způsobu dopravy (STAR EU & CDV, 2014). Scénář je specifikován na cestu 60 km dlouhou a za účelem trávení volného času. Scénáře, které vstupují do analýzy, se mohou od sebe lišit a tato odlišnost může být při analýze zohledněna a dále zkoumána.

Scénář na obrázku obsahuje tři alternativy. Tyto alternativy jsou popsány svými vlastnostmi: dopravním prostředkem, cestovním časem, náklady na cestu a dalšími aspekty. Těmto vlastnostem alternativ se říká atributy. Lišit se od sebe mohou také jednotliví aktéři rozhodování. I jejich například sociodemografické charakteristiky můžeme rovněž zahrnout do analýzy.

## 1.4 Data pro analýzu diskrétních voleb

Do analýzy diskrétních voleb vstupují v zásadě data ze dvou zdrojů. Prvním zdrojem jsou data o skutečných rozhodnutích (tzv. projevených preferencích, *revealed preference* – RP) a druhým zdrojem je experiment, který respondentům situaci pro rozhodnutí simuluje.

Pomocí tohoto experimentu respondenti své preference vyjadřují a hovoříme o vyjádřených preferencích (*stated preference – SP*).

Data o projevených preferencích (*RP data*) jsou data o skutečných rozhodnutích, které lidé udělali. V dopravě to mohou být například volba dopravního prostředku pro určitou cestu, která se zjišťuje pomocí průzkumů dopravního chování (Kouřil et al., 2014). V rámci průzkumu dopravního chování jsou respondenti dotázáni na všechny cesty v rozhodný den. Tato data jsou tedy získána dotazováním na uplynulé chování. V jiných případech mohou být taková data získána přímo pozorováním. Výhodou *RP dat* je, že se týkají skutečných rozhodnutí, které aktéři udělali. Nevýhodou je naopak to, že máme k dispozici pouze údaje o alternativě, kterou aktér nakonec zvolil a chybí nám znalost o nevyužitých alternativách. Zároveň pomocí těchto dat nelze přímo dobře určit model rozhodování například pro zatím neexistující způsob dopravy nebo zásadní chybějící spojení. V dopravních aplikacích se tento nedostatek nahrazuje modelováním těchto chybějících alternativ pomocí dopravního modelu. Jde ale jen o kompromisní řešení. Pokud se totiž změní „pole“ alternativ, které mají lidé k dispozici, mohou se změnit parametry rozhodování těchto lidí. Může se tedy stát, že pokud dosud neexistující alternativní způsoby dopravy budou modelovány dopravním modelem na základě průzkumu stávajícího dopravního rozhodování, budou parametry tohoto rozhodování nevhodné.

Data o vyjádřených preferencích (*SP data*) se získávají pomocí *SP* experimentu, ve kterém je respondentovi předložen simulovaný scénář a respondent má simulovat rozhodnutí v rámci tohoto scénáře. Jasnou nevýhodou je tedy to, že *SP* experiment je „jenom jako“, respondent za své rozhodnutí nenese žádnou zodpovědnost. Tím, že je scénář simulovaný, také vzniká problém realističnosti. Na obrázku 2 je například jako jedna z alternativ osobní automobil, který nemusí mít konkrétní respondent k dispozici. V takovém případě je možné v rámci experimentu nerealistickou alternativu nenabízet. *SP* experiment tedy klade velké nároky na pečlivou přípravu experimentu včetně jeho pilotního vyzkoušení. I tak může být čekání na výsledky celkem napínavé. Velkou výhodou je zkušenost experimentátora a dobrá obeznámenost s problematikou, které se *SP* experiment týká. Výhodou *SP* experimentu je naopak skutečnost, že výzkumník má plnou kontrolu nad nabízenými alternativami a může s nimi manipulovat (proto je to experiment) tak, aby to bylo co nejvýhodnější pro následnou analýzu (Carson & Louviere, 2011). Druhou výhodou je možnost zařadit mezi alternativy nějakou, která dosud neexistuje. Tak je ve scénáři na obrázku 13 jako jedna z alternativ zahrnuta „vysokorychlostní železnice“. Výsledky takového průzkumu poodhalí, jaké postavení by mohla mít vysokorychlostní železnice mezi ostatními způsoby dopravy, kdyby byla mezi Prahou a Brnem uvedena do provozu. Například lze v tomto případě určit očekávaný díl cestujících, kteří mohou přejít na rychlodráhu z jiných způsobů dopravy v důsledku zkrácení cestovního, nebo jaká je citlivost těchto cestujících na cenu rychlovlaku.

## 2. Modely rozhodování

Modelů používaných v analýze rozhodování je celá řada. V této kapitole se budeme zabývat jen některými z nich. Kromě základního multinomiálního modelu nastíníme také problematiku modelů s vnořenou strukturou, modelů s agregovanými alternativami a smíšených modelů. Ukážeme si metodu odhadu parametrů modelů a některá kritéria hodnocení kvality modelů. Nakonec budou zmíněny některé softwarové nástroje pro odhad celé škály různých typů modelů diskretních voleb, především balík mlogit pro R a program Biogeme, se kterými má autor praktickou zkušenost.

### 2.1 Základní model rozhodování

Analýza rozhodování pracuje s výše popsaným pojetím volby. Tato volba je charakterizována scénářem. Scénář obsahuje alternativy a alternativy jsou charakterizovány atributy. Každý respondent zvolí vždy jednu z nabízených alternativ.

V rámci volby lze jakékoli alternativě  $j$  a respondentovi  $i$  přidělit ohodnocení  $U_{ij}$ , kterému se říká užitek. Užitek je bezrozměrná veličina a její velikost má význam pouze ve vztahu k velikostem užiteků ostatních alternativ ve scénáři.

$$U_{ij} = V_{ij} + \varepsilon_{ij} \quad (1)$$

Užitek představuje benefit, jaký by měl respondent  $i$  z toho, kdyby si zvolil alternativu  $j$ . Užitek se skládá z deterministické části  $V_{ij}$  a z náhodné komponenty (chyby)  $\varepsilon_{ij}$ . Deterministická část je obvykle lineární kombinací nezávislých proměnných tak, jak to známe z lineární regrese. Proměnné a jejich parametry (váhy) mohou být několika typů podle toho, s čím (respondentem nebo alternativou) jsou proměnná a odhadovaný parametr asociovány:

$$V_{ij} = a_j + \beta x_{ij} + \gamma_j z_i + \delta_j w_{ij} \quad (2)$$

- Parametr asociovaný s alternativou bez proměnné  $a_j$ . Jedná se o tzv. alternativně specifickou konstantu (*ASC – Alternative Specific Constant*), která je pochopitelně jen jedna pro každou alternativu  $j$  a lze ji interpretovat tak, že respondenti obecně alternativu  $j$  preferují relativně vůči ostatním alternativám. V lineární regresi je parametr asociovaný s alternativou ekvivalentní pojmu intercept.
- Parametr  $\beta$  je stejný pro všechny alternativy a proměnná  $x_{ij}$  je asociovaná s alternativou a popřípadě respondentem (každý respondent může dostat jiné alternativy). To může být například cena, která má ve všech případech stejnou váhu při rozhodování, ale její velikost se liší od alternativy  $k$  alternativě.
- Parametr  $\gamma_j$  je asociovaný s alternativou a proměnná  $z_i$  s individuem. Dejme tomu sociodemografie respondenta tak vstupuje do rovnic všech alternativ, váha jejího působení však může být pro každou alternativu různá. Mohou se například lišit

preferance různých dopravních módů u lidí různého věku nebo společenského postavení.

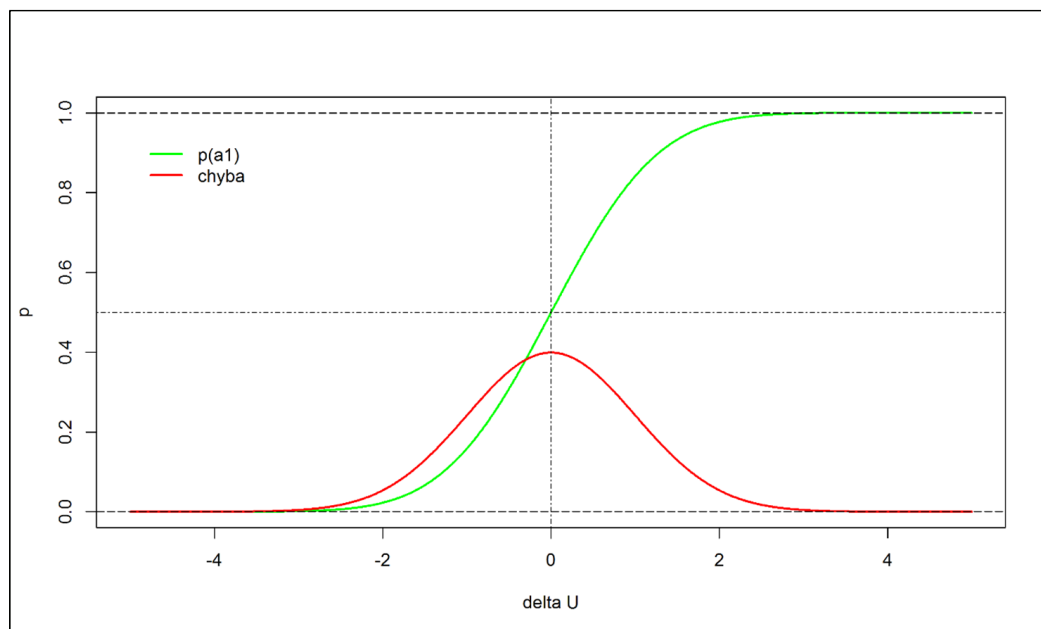
- Parametr  $\delta_j$  i proměnná  $w_{ij}$  jsou asociovány s alternativou. Proměnná opět navíc s individuem, protože každý respondent může dostat k výběru různé alternativy. Například cestovní čas se liší v různých alternativách dopravy, ale zároveň může být jeho váha závislá na konkrétní alternativě. Váha cestovního času (tato váha má negativní užitek) ve veřejné dopravě může být menší, než v dopravě individuální, protože se cestující po dobu cesty veřejnou dopravou může věnovat odpočinku nebo práci.

Velikost parametrů v modelu je odhadována obvykle metodou největší věrohodnosti, která bude probrána ve čtvrté kapitole.

Nyní ještě krátce k náhodné komponentě (chybě) modelu  $\varepsilon_{ij}$ . Manski (1977) identifikoval čtyři možné zdroje náhodné komponenty. Mohou jimi být nepozorované vlastnosti alternativ, nepozorované individuální vlastnosti, chyby měření nebo vlastnosti měřícího nástroje. Nepozorovanými vlastnostmi alternativ se myslí atributy alternativ, které model nezachycuje. Například při volbě dopravního módu může hrát roli, jestli lze příslušným dopravním prostředkem přepravovat větší zavazadlo (například lyže). Obvykle se taková možnost v modelech nemodeluje, ale v určitých situacích může být důležitá. Budeme-li se například zabývat využíváním veřejné dopravy v horských rekreačních oblastech, budeme patrně s touto vlastností dopravní alternativy počítat. Nepozorovanými individuálními vlastnostmi se myslí takové vlastnosti jednotlivých aktérů rozhodování, které nebyly zachyceny v modelu jako individuálně závislé proměnné. Například někteří cestující potřebují najít místo k sezení, jiní s sebou přepravují těžké zavazadlo a podobně. Tyto nepozorované individuální rozdíly mohou ovlivňovat individuální hodnoty parametrů. Tím, že cílem odhadu je stanovení parametrů obecně pro všechny individua, vstoupí tyto nepozorované rozdíly mezi individui do chyby modelu.

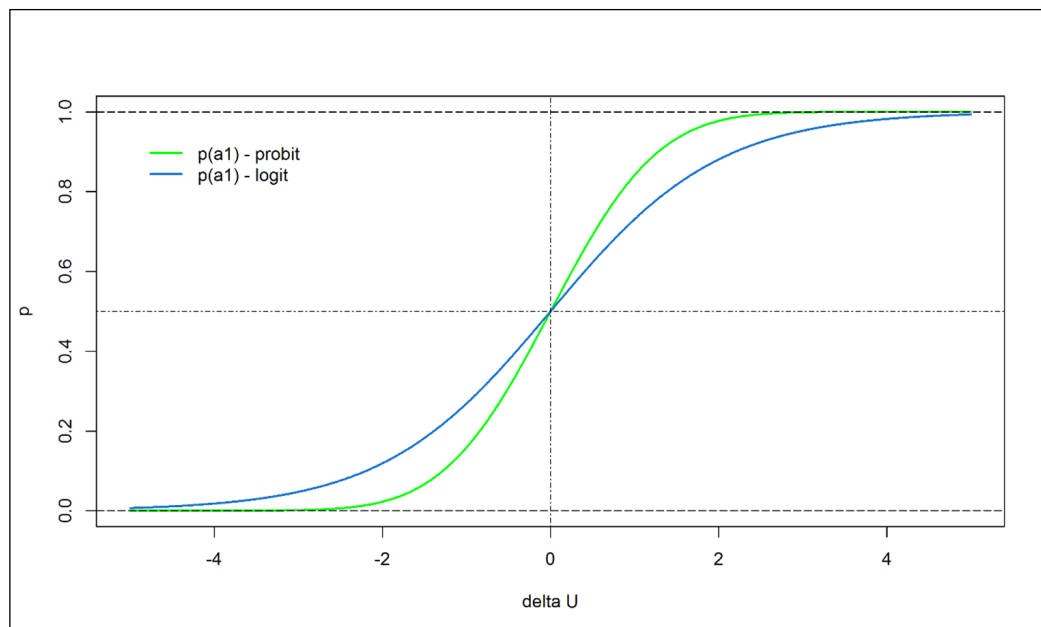
Model předpokládá, že bude vybrána alternativa s největším užitekem. Přítomnost náhodné komponenty  $\varepsilon_{in}$  ovšem znamená, že nestačí jednoduše vypočítat velikosti deterministických částí užitkových funkcí a potom vzít alternativu s největším užitekem. Mezi užitekem z určité alternativy a pravděpodobností jejího zvolení se proto nachází link (převodní) funkce. Čím bude užitek nějaké alternativy větší než užitky ostatních alternativ, tím poroste pravděpodobnost jejího zvolení. Nikdy však tato pravděpodobnost nebude absolutní jistota, protože u chybové funkce  $\varepsilon_{in}$  se předpokládá Gaussovské rozdělení, které pro všechny reálné hodnoty chyby nabývá nenulové pravděpodobnosti výskytu. Nabízí se funkce kumulativního normálního rozdělení (probitová funkce).

Představme si příklad scénáře se dvěma alternativami ( $a_1$ ,  $a_2$ ). Potom v závislosti na chybě bude křivka, která vyjadřuje pravděpodobnost volby alternativy  $a_1$ , záviset na rozdílném užtku ( $\Delta U$ ) a velikosti chyby.



Obrázek 3: Křivka probitové funkce a normální rozdělení chyby

V praxi se častěji než modely, založené na probitové funkci, používají modely vycházející z logitové funkce.



Obrázek 4: Křivky probitové vs. logitové funkce



Logitová funkce má pozvolnější sklon, než funkce probitová a na krajích není tolik plochá, což usnadňuje odhad modelu. Logitová funkce pro dvojici alternativ (binary logit) má následující matematický tvar.

$$p_1 = \frac{e^{U_1 - U_2}}{1 + e^{U_1 - U_2}} \quad (3)$$

Kde:

$p_1$  je pravděpodobnost zvolení první z alternativ a

$U_1 - U_2$  je rozdíl mezi užitky první a druhé alternativy

Křivky na obrázku 3 a 4, stejně jako rovnice logitové funkce, reprezentují pouze případy rozhodování se dvěma alternativami. Modely s více alternativami (*multinomial logit* – *MNL*) používají jako základní převod mezi velikostmi užitku jednotlivých alternativ a pravděpodobnostmi jejich zvolení nejčastěji funkci, která je zobecněním logitové funkce pro více alternativ.

$$p_{ij} = \frac{e^{U_{ij}}}{\sum_{k=1}^M e^{U_{ik}}} \quad (4)$$

Kde:

$p_{ij}$  je pravděpodobnost výběru alternativy  $j$  respondentem  $i$ ,

$U_{ij}, U_{ik}$  je užitek z výběru alternativy  $j$  respektive  $k$  respondentem  $i$ ,

$M$  je počet všech alternativ ve scénáři.

Pravděpodobnost zvolení alternativy se tedy počítá jako exponenciála jejího užitku, která je normována přes všechny alternativy tak, aby výsledný součet pravděpodobností zvolení všech alternativ byl roven jedné.

Jako příklad si představme scénář, ve kterém jsou tři alternativy. Užitek první z alternativ je roven 1 (ať už to pro tuto chvíli znamená cokoli), užitek druhé alternativy je dvakrát větší a užitek alternativy třetí je větší třikrát. V následující tabulce jsou ve sloupcích užitky alternativ  $U$ , jejich exponenciální hodnoty

$e^U$  a výsledné pravděpodobnosti zvolení  $p$ .

Tabulka 1: Příklad hodnoty užitku a pravděpodobnosti výběru každé ze tří alternativ

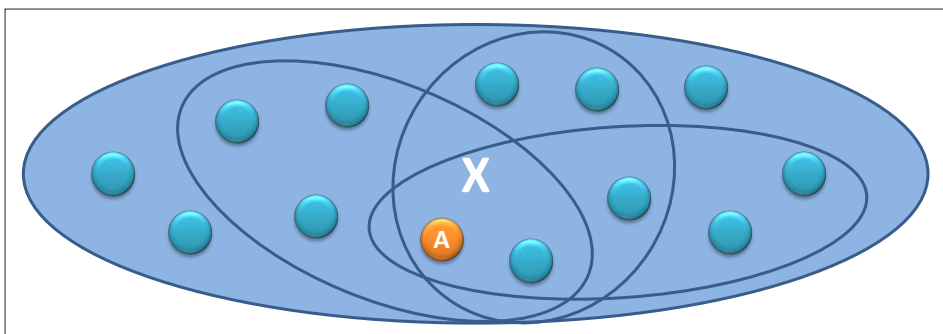
Alternativa	$U$	$e^U$	$p$
1.	1	2,718	0,090
2.	2	7,389	0,245
3.	3	20,086	0,665

Pravděpodobnosti zvolení jednotlivých alternativ jsou menší než 1 a součet těchto pravděpodobností je roven jedné. Největší pravděpodobnost zvolení má, podle očekávání,

třetí alternativa, která, ačkoli je její užitek jen 3x větší než užitek alternativy první a 1,5x větší než užitek druhé alternativy, bude volena s více než 7x větší pravděpodobností, než první alternativa a zhruba s 3x větší pravděpodobností, než alternativa druhá. Z rovnice 4 vyplývá, že pravděpodobnost výběru nějaké alternativy závisí nejen na jejím užtku, ale také na užtících ostatních alternativ.

## 2.2 Princip nezávislosti irelevantních alternativ

Představme si situaci rozhodování, kdy je celkově k dispozici množina  $X$  alternativ. Z této množiny decision maker vybírá alternativu  $A$ . Alternativa  $A$  má pro rozhodčího nejvyšší užitečnost. Proto byla vybrána z celé množiny alternativ  $X$ . Intuitivně je zřejmé, že pokud vybereme jakoukoli podmnožinu z množiny alternativ  $X$ , která zároveň obsahuje alternativu  $A$ , vybere tento rozhodčí vždy opět alternativu  $A$ .



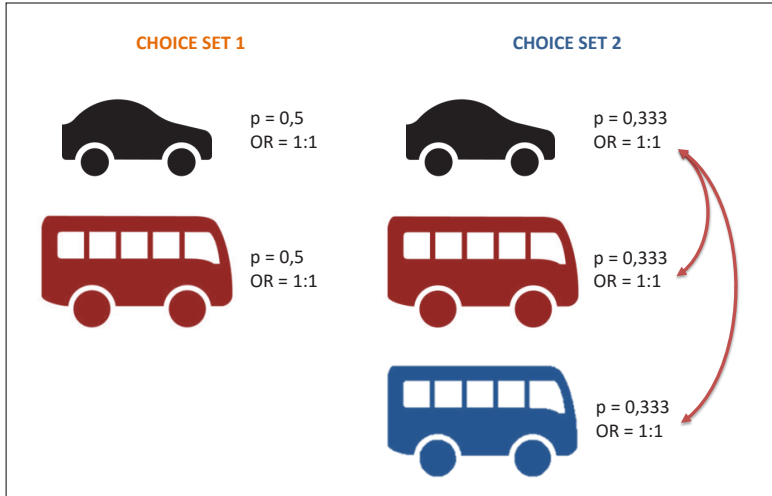
Obrázek 5: Varianty s preferovanou alternativou a různým počtem ostatních alternativ

Výběr je tedy nezávislý na méně preferovaných alternativách, které jsou do množiny výběru zařazeny. Ať už jsou v množině, ze které se vybírá jakékoli méně preferované alternativy a v jakémkoli množství, vybere rozhodčí vždy nejvíce preferovanou alternativu  $A$ . Tomu se říká princip nezávislosti irelevantních alternativ (*Independence of Irrelevant Alternatives – IIA*). Například, pokud právě nejvíce ze všeho preferujete trávit čas v přírodě, měli byste si vybrat tuto alternativu bez ohledu na ostatní alternativy (dejme tomu domov a centrum města) a pro přírodu byste se měli rozhodnout také, pokud by se do výběru alternativ jako další místo přidala obchodní galerie.

Zatímco v případě rozhodnutí jednotlivce je princip nezávislosti irelevantních alternativ intuitivně dobře pochopitelný, podívejme se, jaké má tento princip důsledky pro pravděpodobnosti rozhodování většího souboru rozhodujících. V takovém souboru už hraje roli náhodná komponenta modelu, a to individuální rozdíly v preferencích. Neexistuje jen jedna alternativa, která je pro všechny nejlákavější.

Nyní si představme nějakou trasu, po které se lidé mohou dopravovat osobním automobilem nebo autobusem. Výběr automobilu nebo autobusu vychází půl na půl. Pravděpodobnost výběru automobilu (nebo autobusu) je mezi lidmi 0,5 (tedy 50 %), poměr šancí výběru automobilu (autobusu) je 1:1. Poměr šancí se ve statistice nazývá anglicky *odds ratio*

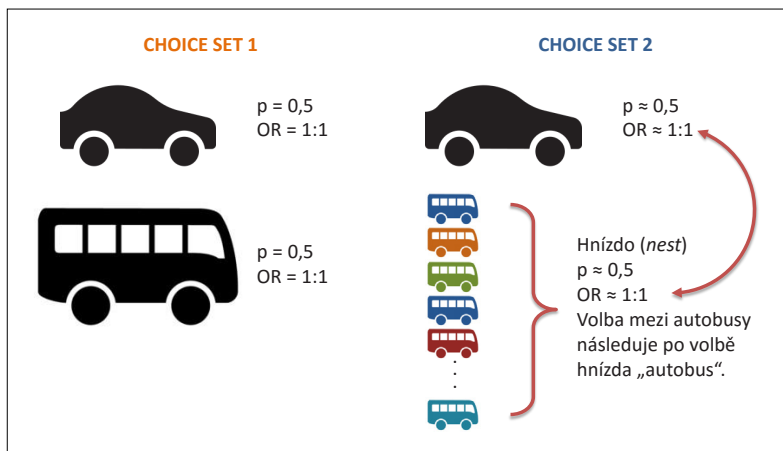
(OR) a představuje důležitou míru, jak uvidíme dále. Situaci výběru mezi automobilem a autobusem odpovídá *choice set 1* na obrázku 6.



Obrázek 6: Problém červeného a modrého autobusu

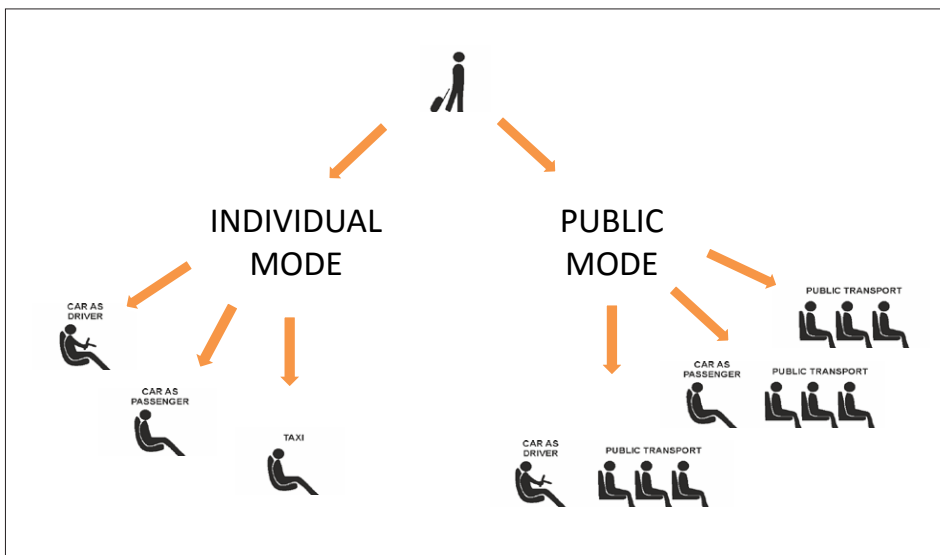
### 2.3 Modely s vnořenou strukturou – nested logit

Představme si nyní, že místo červeného a modrého autobusu je na trase zřízena celá linka nebo několik linek. Cestující si nyní může vybrat mezi osobním automobilem a různými autobusovými spoji. Tyto různé autobusové spoje mají poněkud odlišné jízdní doby a i ceny za přepravu se liší. Těžko předpokládat, že cestující bude vážit mezi osobním automobilem a každým z nabízených autobusových spojů. Dá se očekávat, že nejprve rozhodne mezi cestování automobilem a nějakým autobusem. Teprve poté si cestující vybere konkrétní autobusový spoj, například podle času odjezdu.



Obrázek 7: Vnořená struktura

Rozhodování tedy probíhá ve dvou krocích. Konkrétní autobusový spoj je vybrán v procesu rozhodování, které je vnořeno do rozhodování mezi automobilem a některým z autobusových spojů. Vnořené struktury se říká *nest* (hnízd) a model, který jej popisuje, se nazývá *nested logit*. Pro vnořenou strukturu jsou v modelu odhadovány navíc škálovací parametry pro lepší odhad ostatních parametrů sobě podobných alternativ (Hensher et al., 2005).



Obrázek 8: Vnořená struktura volby dopravního módu cestování na letiště (cdv & median, 2017)

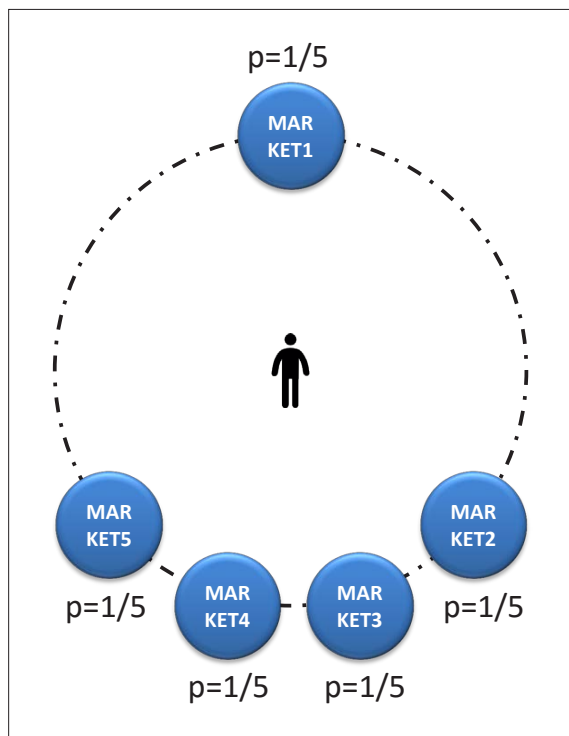
*Nested logit* modely se používají často při analýzách volby dopravního módu. Někdy se ukazuje jako užitečné rozlišení na motorizované a nemotorizované módy nebo na individuální a veřejnou dopravu. Ve studii, kde jsme zjišťovali způsob rozhodování leteckých cestujících při cestování na letiště Praha (CDV & MEDIAN, 2017). Jako vhodné se ukázalo dát dohromady takové způsoby dopravy na letiště, které obsahovaly veřejnou dopravu a zvláště nechat čistě individuální dopravu na letiště (obrázek 8). Vnořenou strukturu při analýze volby dopravního módu použijeme v případové studii volby dopravního prostředku na cestu po brněnské metropolitní oblasti (kapitola 6).

Vnořená struktura může mít víc úrovní zanoření, nebo může reflektovat další korelační vztahy mezi členy různých hnízd (*cross nested*). Modely pro tyto složitější případy jsou popsány v řadě publikací (Ben-Akiva & Lerman, 1985).

## 2.4 Model pro agregované alternativy

Někdy je nutné modelovat tak, že některá z alternativ ve skutečnosti představuje souhrn více elementárních alternativ. Situace je podobná jako na předchozím případě s vnořenou strukturou, kde volba autobusové dopravy v sobě zahrnovala možnost cestovat jedním z několika autobusových spojů.

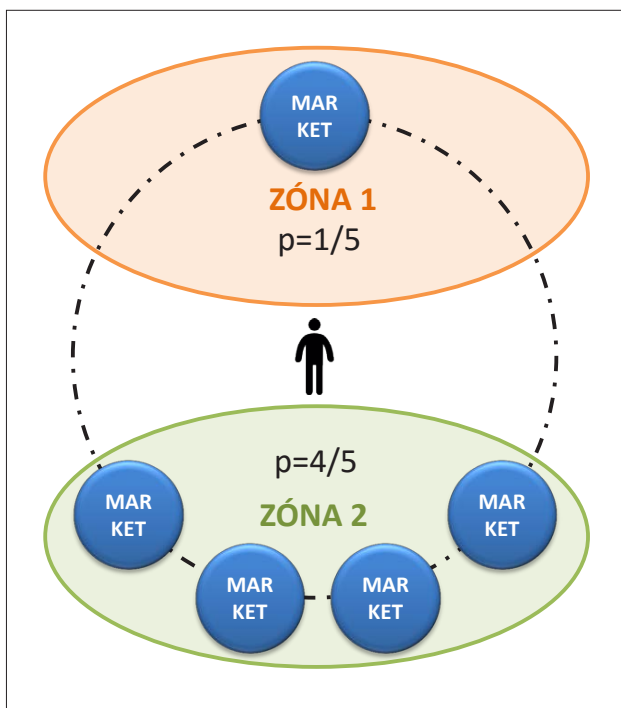
Uvažujme ilustrativní situaci, kdy se rozhodujeme, kam si zajdeme nakoupit. Do úvahy přichází pět obchodů, které jsou ve všech ohledech stejné a také jsou od nás stejně vzdálené – jde o modelový příklad. Předpokládejme, že v každém případě stačí navštívit pouze jeden z těchto obchodů. Situace je znázorněna na obrázku.



Obrázek 9: Pravděpodobnost výběru jednoho z pěti stejných obchodů k nákupu

Intuitivně je zřejmé, že jsou-li všechny tyto obchody stejné a stejně vzdálené, bude pro nás užitek z nákupu v kterémkoli z nich rovněž stejný a také pravděpodobnost výběru kteréhokoli z těchto pěti stejných obchodů k nákupu bude stejná, a to  $1/5$ . Kdybychom nějak chtěli vyjádřit užitek z návštěvy každého jednoho z obchodů, můžeme klidně obchodům přidělit užitek v hodnotě 1. Užitek je bezrozměrné číslo a jediné, co musíme dodržet je, aby byl pro každý z obchodů stejný. Navíc by měl být kladný, protože přece nějaký obchod navštívit chceme. Jeden obchod se bude rovnat jednomu dílu užitku.

Nyní si ale představme, že území je rozděleno do dvou zón (tak tomu bývá například v dopravních modelech). V první zóně je jeden z těchto pěti obchodů, zatímco ve druhé zóně jsou zbývající čtyři obchody. Otázka nyní je, s jakou pravděpodobností zvolíme kterou z těchto dvou zón k realizování nákupu.



Obrázek 10: Pravděpodobnost výběru jedné ze dvou zón k nákupu

Intuice nám opět napovídá, že tyto dvě zóny se na rozdíl od samotných obchodů liší. V předchozím zjištění jsme přišli na to, že každý z obchodů navštívíme se stejnou pravděpodobností  $1/5$ . Z toho vyplývá, že do zóny 1 půjdeme s pravděpodobností  $1/5$ , zatímco do zóny 2 s pravděpodobností  $4/5$  (jsou zde 4 obchody). Abychom mohli takovou situaci modelovat pomocí modelu diskretních voleb, musíme nějak spočítat užitekovou funkci pro obě zóny. Je zřejmé, že užitek z návštěvy zóny 1 nebo 2 nebude stejný, zároveň ale neplatí, že by měl být užitek z návštěvy zóny 2 čtyřikrát větší, než užitek z návštěvy zóny 1. Nestačí prostě sečíst užítky obchodů uvnitř jednotlivých zón. Je to tak proto, že našim cílem je navštívit pouze jeden z obchodů, ať jich je v zóně kolik chce. Řešení pro výpočet užitku z návštěvy zóny 2 je ale jen o málo složitější. Je-li užitek z návštěvy zóny 1 roven jedné (jeden obchod), potom užitek z návštěvy zóny 2 je přirozený logaritmus počtu obchodů (tedy čtyř). Ve skutečnosti je užitek z návštěvy první zóny rovněž logaritmus z počtu obchodů v této zóně, logaritmus jedné je roven zase 1.

$$U_{zóna 1} = \ln(1) = 1 \quad (5)$$

$$U_{zóna 2} = \ln(4) = 1,386294$$

Když takto vypočítané hodnoty užitekových funkcí dosadíme do vzorce pro výpočet pravděpodobnosti (3) zjistíme, že dostáváme takové pravděpodobnosti návštěvy obou zón, které jsme předpokládali, a to  $1/5$  a  $4/5$ .

Určité zvýšení složitosti modelu představují méně idealistické scénáře, ve kterých například obchody nejsou stejné. V takovém případě nejprve zjistíme průměrnou míru užítku všech  $M_a$  elementárních alternativ, které agregovaná alternativa  $a$  zahrnuje. Je to průměr  $\bar{v}_{ia}$  užítků všech  $M_a$  elementárních alternativ v agregované alternativě  $a$ .

$$\bar{v}_a = \frac{1}{M_a} \sum_{i=1}^{M_a} v_{ia} \quad (6)$$

Následně vypočítáme míru heterogenity všech  $M_a$  elementárních alternativ, které zahrnuje agregovaná alternativa  $a$ .

$$B_a = \frac{1}{M_a} \sum_{i=1}^{M_a} e^{v_{ia} - \bar{v}_a} \quad (7)$$

Užitek agregované alternativy  $V_a$  pak dostaneme sečtením průměrného užítku elementárních alternativ  $\bar{v}_a$  s přirozeným logaritmem jejich počtu  $M_a$  a přirozeným logaritmem míry variability  $B_a$ .

$$V_a = \bar{v}_a + \ln(M_a) + \ln(B_a) \quad (8)$$

Ve většině případů, kdy potřebujeme použít agregované alternativy, nedokážeme přímo vypočítat užítky jejich elementárních alternativ a ani nemusíme znát ani jejich počet. V takových případech se průměrný užitek  $\bar{v}_a$  elementárních alternativ a v agregované alternativě nahradí lineární regresí proměnných, které ukazují na užitek těchto elementárních alternativ, počet elementárních alternativ  $M_a$  nahradí lineární regresí ukazatelů na „rozlohu“ agregované alternativy  $a$  a míra heterogenity  $B_a$  se nahradí lineární regresí ukazatelů variability elementárních alternativ v agregované alternativě  $a$ . Odvození těchto vzorců a další možnosti práce s agregovanými alternativami jsou popsány v reportu (Ben-Akiva, Gunn & Silman, 1984) anebo v knize (Ben-Akiva & Lerman, 1985).

## 2.5 Smíšené modely

Stejně jako u jiných regresních modelů lze i v modelech rozhodování použít náhodné koeficienty. Obvykle se této techniky využívá při odhadu modelů na tzv. panelových datech.

Panelová data obsahují více pozorování (pozorovaných rozhodnutí) od každého respondenta. V praxi totiž bývá obtížné (drahé) získat respondenta ke spolupráci, hodně času zabere také získání dalších potřebných údajů od respondenta, ale samotné provedení *SP* experimentu je otázka chvilky. Z tohoto důvodu bývají respondenti podrobni několika různým scénářům *SP* experimentu. Když ale titíž respondenti provádí několik *SP* experimentů, může u nich existovat individuální efekt, který se projeví ve všech pokusech, které respondent absolvoval. Protože rozhodování od jednoho respondenta na sobě nejsou nezávislá, není splněna

podmínka nezávislosti jednotlivých pozorování na sobě navzájem. To ohrožuje korektní odhad parametrů modelu. Situace se řeší tím, že se specifika jednotlivých respondentů zahrnou do modelu v podobě variance odhadovaných parametrů (Hensher & Greene, 2001). Do modelu se zavádí místo obvyklých odhadovaných parametrů modelu náhodné proměnné, u kterých se předpokládá určitá variabilita napříč respondenty. Odhaduje se pak průměrná velikost takového parametru i jeho variabilita. Smíšený (*mixed*) model pro panelová data bude použit v této monografii v rámci případové studie volby studentského pronájmu (kapitola 6).

## 2.6 Software pro analýzu diskrétních voleb

Nástroje pro analýzu diskrétních voleb jsou obsaženy snad ve všech komerčních statistických balíčcích (*Stata, Statistica, SPSS...*). Významným komerčním nástrojem je také Nlogit (Greene, 2002; Hensher et al., 2007), přičemž autoři a propagátoři tohoto nástroje patří k předním výzkumníkům analýzy rozhodování.

Volně dostupným nástrojem je software *Biogeme* (Bierlaire, 2003). *Biogeme* je vyvíjen na École polytechnique fédérale de Lausanne - Epfl ve Švýcarsku<sup>1</sup>. Program je zde ke stažení v několika verzích a dále též řada dokumentů a příkladů. *Biogeme* je vhodný pro formulaci modelů s tzv. pojmenovanými alternativami (viz kapitola 3.1). Jeho výhodou je snadná formulace odlišných užitekových funkcí pro každou alternativu včetně nelinearit v modelu. Tento nástroj je dobře použitelný například k přípravě modelu volby dopravního módu (*mode choice*) dopravních modelů. Další verze *Biogeme* je vytvářena jako extenze programovacího jazyka Python (Bierlaire, 2016), čímž za cenu přechodu na jinou syntaxi získává rozšiřující možnosti přípravy dat, analýz a prezentací výsledků, které Python umožňuje.

Prostředí statistického nástroje R (R Core Team, 2018) nabízí mimo jiné balík *mlogit* (Croissant, 2011, 2012; Viton, 2014), který se hodí především pro odhad modelů s nepojmenovanými alternativami. Přímo neumožňuje psát odlišné užitekové funkce pro jednotlivé alternativy, zato nabízí více jiných možností nastavení a komplexní prostředí R pro úpravu dat včetně přípravy optimálního designu (kapitola 4.2), nebo prezentaci výsledků. Balík *mlogit* se hodí zejména pro exploraci a testování hypotéz a pro vytváření symetrických modelů, kde mají všechny alternativy stejné rovnice užitekových funkcí. Pokud se ovšem smíříme s trochou programování, lze *mlogit* označit za vhodný pro všechny obvyklé případy. Z tohoto důvodu budeme praktické postupy a kazuistické příklady v této knize prezentovat právě v prostředí R a s balíkem *mlogit*.

Novějším nástrojem pro analýzu multinomiálních dat v prostředí R je balík *gmnl* (Sarrias, 2016; Sarrias & Daziano, 2017). Tento balík volí podobný přístup k datům jako balík *mlogit*. Také způsob ovládání je velmi podobný, pokud hodláte dělat stejné věci. Na rozdíl od balíčku *mlogit* umožňuje *gmnl* navíc odhad latentních tříd (Lenk, 2011).

1 <http://biogeme.epfl.ch/>

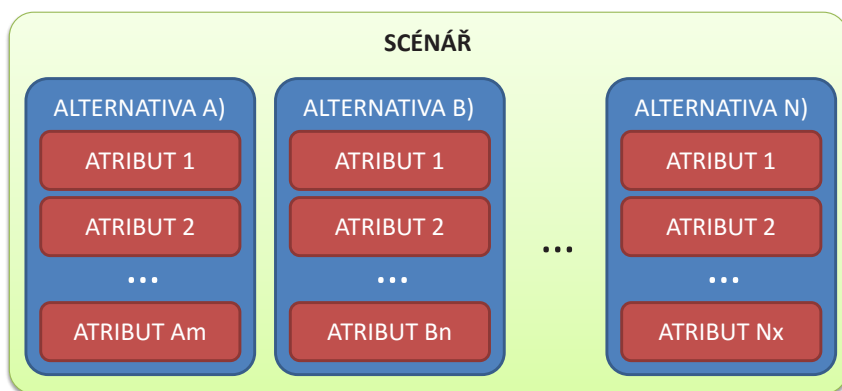


### 3. Tvorba experimentu

V této kapitole se budeme zabývat tvorbou *SP* experimentu. Podíváme se podrobněji, z čeho se skládá scénář a jaká jsou doporučení pro jeho vytváření. Po té se budeme zabývat tím, jak sestavit jednotlivé varianty scénářů a jak je efektivně namíchat. Budeme se zabývat personalizací experimentu a adaptivními postupy. Krátce probereme způsoby výběru respondentů a zmíníme význam pilotní studie.

#### 3.1 Příprava scénáře

Jak již bylo popsáno výše, scénář se skládá z alternativ a alternativy jsou popsány pomocí atributů. Alternativy musí být minimálně dvě (jinak by nešlo o rozhodování) a v případě *SP* experimentu jich má být maximálně jen takové množství, aby jimi respondent nebyl zahlcen. Naprostá většina scénářů mívá do čtyř alternativ.



Obrázek 11: Struktura rozhodování v *SP* experimentu

Alternativy mohou být takzvaně pojmenované a nepojmenované. Například při volbě dopravního módu se často ve scénáři vyskytují alternativy, které zastupují jednotlivé způsoby dopravy (např. obr. 2). Alternativy jsou pojmenované podle těchto dopravních módů a mohou mít různé atributy. Například pěší nebo cyklistický mód nemívá atribut ceny za dopravu. Při odhadu parametrů se pak odhadují pro každou alternativu alternativně specifické konstanty. Scénář s nepojmenovanými alternativami je naopak na obrázku 11. Pro takový design se alternativně specifické konstanty neodhadují. Oba přístupy jsou do značné míry zaměnitelné a jeden nebo druhý se používají podle zvyklosti. V modelu s nepojmenovanými alternativami lze například zapsat pojmenované alternativy tak, že se do modelu zavede nominální proměnná jména alternativy. Parametry této proměnné pak v modelu s nepojmenovanými alternativami zastupují alternativně specifické konstanty.

Scénáře mají podle doporučení obsahovat maximum realistických alternativ, které jsou v současnosti k dispozici. Atributy a jejich úrovně mají být definovány čistě a jednoznačně, dobře srozumitelné pro respondenty a také interpretovatelné pro výzkumníky. Rovněž jsou

doporučovány scénáře, které obsahují „no choice“ volbu (Johnston et al., 2017). Použití možnosti, která respondentovi dovolí nevybrat žádnou z předložených alternativ, nebo rozhodnutí odložit může zlepšit validitu výsledků. Pozorování, která obsahují „no choice“ volbu pak je možné z další analýzy vyloučit, nebo je lze zpracovávat jako alternativu, která kromě alternativně specifické konstanty nemá žádné alternativně specifické atributy (Louviere, Hensher & Swait, 2000b).

Někdy není zcela jasné, co má vliv na rozhodování a co by tedy v podobě atributů mělo figurovat ve scénářích. Za tímto účelem mohou být před samotný *SP* experiment zařazeny kvalitativní studie, například focus groups, které mají sloužit k identifikaci relevantních vlastností alternativ. Atributů u alternativ by mělo být opět jen takové množství, aby respondenta nezahltily a aby se podle nich byl respondent schopen rozhodovat.

### 3.2 Design experimentu

Když je dostatečně dobře definovaný scénář, bylo rozhodnuto o počtu alternativ, jsou předběžně určeny požadované atributy těchto alternativ a jejich úrovně, může se přejít k sestavování jednotlivých scénářů, jak budou předkládány respondentům. Tomu se říká design experimentu. Design experimentu je způsob, jak manipulovat s atributy a jejich úrovněmi tak, aby to co nejlépe umožnilo důkladné testování hypotéz, které jsou předmětem našeho zájmu (Louviere, Hensher & Swait, 2000b: s. 84).

Scénáře předkládané respondentům mají variovat takovým způsobem, aby se zajistila co nejmenší vzájemná korelace atributů mezi scénáři. Čím těsněji spolu dva atributy v jakémkoli regresním modelu korelují, tím obtížnější je odlišit jejich individuální účinek na závislou proměnnou. Nejlepší je, pokud spolu atributy nekorelují vůbec. Ideálem je zde tzv. faktoriální design, kdy jsou scénáře složeny tak, že obsahují všechny možné kombinace úrovní atributů.

Tabulka 2: Faktoriální design pro hlavní přísady v polévce (Louviere et al., 2000)

Kombinace	Maso	Nudle	Zelenina
1	kuře	ano	ano
2	kuře	ano	ne
3	kuře	ne	ano
4	kuře	ne	ne
5	hovězí	ano	ano
6	hovězí	ano	ne
7	hovězí	ne	ano
8	hovězí	ne	ne

V tabulce je ukázka faktoriálního designu, kde jsou tři atributy a všechny tři dichotomické. Jejich kombinace dávají dohromady 8 možností.

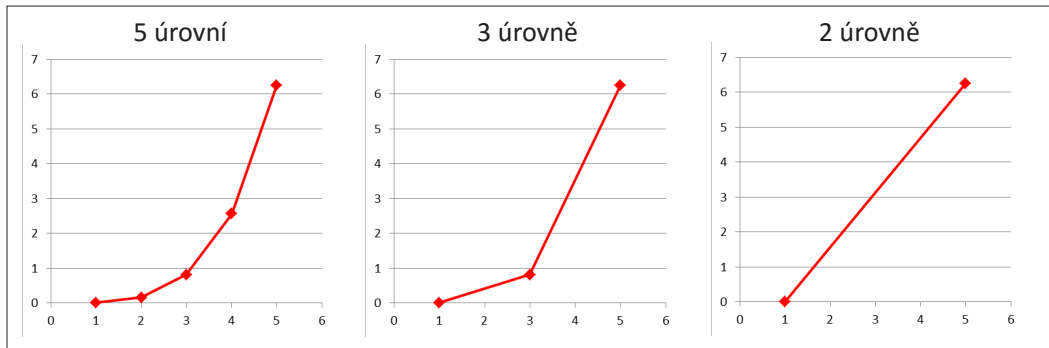
Faktoriální design zaručuje kompletní vzájemnou nezávislost všech atributů. Takový design teoreticky umožňuje studovat všechny vztahy mezi atributy, a to včetně jejich interakcí.

Počet kombinací narůstá s počtem atributů a jejich úrovní. V případě, že experiment pracuje s pojmenovanými alternativami také s počtem alternativ. Počet kombinací a tedy scénářů s  $M$  pojmenovanými alternativami,  $A$  atributy v každé alternativě a  $L$  úrovněmi atributů získáme podle následujícího vzorce.

$$n_{M,A,L} = L^{MA} \quad (9)$$

Počet různých scénářů ve faktoriálním designu velmi rychle narůstá se zvětšením kterékoli z uvedených proměnných.

Počet úrovní atributů také zvyšuje počet scénářů v plně faktoriálním designu. Stojí tedy za úvahu množství úrovní atributů redukovat. Méně úrovní atributu může opomíjet důležité nelinearity, ale dvě úrovně stačí k postižení lineárního trendu. Pokud tedy neplánujeme zkoumat nelineární působení atributu, budou pouhé dvě úrovně stačit. Pokud je působení nějakého atributu nelineární, bude výsledek záviset na intervalu mezi zvolenými dvěma úrovněmi. Podle toho dostaneme sílu působení.



Obrázek 12: Úrovně nelineárního atributu

Jinou možností je redukovat plně faktoriální design tak, aby spolu nekorelovaly ty atributy nebo jejich kombinace, které nás zajímají. K tomu se používá tzv. ortogonální kódování atributů. Hodnoty (úrovně) atributů se zakódují tak, aby součet všech těchto úrovní atributů byl nulový. Pokud má atribut dvě úrovně, bude se první úroveň kódovat -1 a druhá 1. Pokud má atribut tři úrovně, bude se první úroveň kódovat -1, druhá 0 a třetí 1. A tak dále podle tabulky.

Tabulka 3: Ortogonální kódování úrovní atributů

Počet úrovní atributu	Úroveň atributu				
	1.	2.	3.	4.	5.
2	-1	1			
3	-1	0	1		
4	-2	-1	1	2	
5	-2	-1	0	1	2

Nyní můžeme překódovat ukázkou kombinací přísad v polévce do ortogonálního kódování. Protože atributy mají všechny jenom dvě úrovně, můžeme například hodnotou -1 kódovat „kuře“ a „ne“ a hodnotou 1 zase „hovězí“ a „ano“. Zároveň také se změnou kódování opustíme zcela příklad s polévkou a atributy nazveme jednoduše  $A$ ,  $B$  a  $C$ .

Tabulka 4: Ortogonální kódování tří atributů a jejich kombinací

Kombinace	A	B	C	$A*B=D$	$A*C=E$	$B*C=F$	$A*B*C=G$
1	-1	1	1	-1	-1	1	-1
2	-1	1	-1	-1	1	-1	1
3	-1	-1	1	1	-1	-1	1
4	-1	-1	-1	1	1	1	-1
5	1	1	1	1	1	1	1
6	1	1	-1	1	-1	-1	-1
7	1	-1	1	-1	1	-1	-1
8	1	-1	-1	-1	-1	1	1

Když bychom nyní korelovali ortogonálně kódované atributy ( $A$ ,  $B$ ,  $C$ ) mezi sebou, byly by všechny korelace nulové, protože atributy jsou na sobě zcela nezávislé. Stejně by byly nulové korelace kombinací atributů mezi sebou ( $A*B$ ,  $A*C$ ,  $B*C$ ) nebo s čistými atributy i kombinace všech tří atributů ( $A*B*C$ ) se zbytkem tabulky.

Dejme tomu, že nás zajímají pouze některé efekty. V tom případě, z množiny plně faktoriálního designu vybereme takovou podmnožinu scénářů, ve které nebudou korelovat pouze ty sloupce, které zastupují efekty, které nás zajímají. Takovýmto dílčím výběrem z faktoriálního designu se říká optimální designy a jsou podrobněji popsány například v knize Louviera et al. (2000). V prostředí R lze ke generování designů použít balíček *AlgDesign* (Wheeler & Braun, 2015).

Pokud nás zajímají pouze hlavní efekty, můžeme k problému přistoupit z druhé strany. Víme, že v tabulce 4 jsou všechny atributy ve sloupcích nekorelované (je to faktoriální design), ale zároveň jsou nekorelované také společně se čtyřmi jejich kombinacemi. Je-li našim požadavkem pouze aby nekorelovaly samotné atributy, můžeme čtyři kombinované sloupce ( $A*B$ ,  $A*C$ ,  $B*C$ ,  $A*B*C$ ) prohlásit za atributy  $D$ ,  $E$ ,  $F$  a  $G$ . Získali jsme tak takzvaný *frakcionální design*, kde je v osmi kombinacích až sedm nekorelovaných atributů. Tento design ovšem nebude schopen v žádném případě podchytit jakékoli interakce mezi atributy. Pokud použijeme pouze *frakcionální design* například s atributy  $A$ ,  $B$ ,  $C$  a  $D$ , nedokáže tento frakcionální design podchytit interakce atributů  $A$  a  $B$  a odlišit ji od účinku atributu  $D$ .

Experimentální design je efektivní nástroj, který nám pomůže navrhnout takové schéma testování, které při relativně velmi malém vzorku respondentů povede k dobrým odhadům předpokládaného modelu. Hodí se tedy tam, kde dopředu víme, jaké modely budeme odhadovat a dokážeme respondenty předem přesně zařadit.

### 3.3 Personalizace experimentu

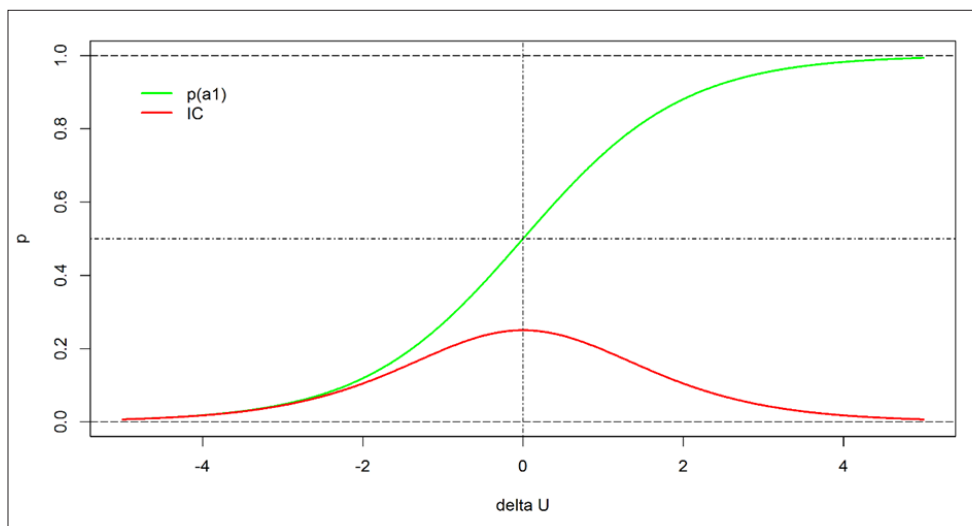
Experiment by měl pracovat se scénářem, který je z hlediska respondenta realistický. Proto se používají postupy, které scénáře upravují respondentovi na míru. V rámci designu se tak například zajišťuje, aby osoby, které nejsou výdělečně činné, nerozhodovaly o volbě dopravního módu na cestě do zaměstnání. Respondentům, kteří nemají k dispozici nějaký způsob dopravy (například automobil) není mezi alternativami tento způsob dopravy nabízen, a podobně.

Další možností je napojení *SP* experimentu na skutečné rozhodnutí, které respondent realizoval. *SP* experiment může například vycházet z konkrétní cesty, kterou respondent absolvoval v minulosti, přičemž se v takovém případě můžeme doptat na charakteristiky cesty (například cestovní čas). Ostatní alternativy však budeme muset do experimentu namodelovat tak, aby byly co nejvíc v souladu s předpokládanou žitou zkušeností respondenta. Takový postup je velmi náročný na přípravu, a často klade i značné nároky na kvalitu práce tazatele a neobejde se bez počítačem asistovaného dotazování. Na druhou stranu vede podle předpokladů k nejvíce realistickým výsledkům, protože nejlépe odpovídá respondentově realitě.

### 3.4 Adaptivní testování

V klasickém (neadaptivním) přístupu se nastaví změny atributů v experimentálním designu předem. Výzkumník postupuje podle zkušenosti a nejlépe s využitím pilotní studie. Při takovém přístupu hrozí, že se předem nastavené změny atributů budou lišit málo a výsledky experimentu neodhalí změny v rozhodování na základě změn těchto atributů. Rovněž je zde riziko příliš velkých rozdílů ve změnách atributů. V takovém případě sice bude ke změnám chování na základě změn atributů docházet, ale nebude z toho možné zjistit kvůli velkým rozdílům v attributech nové informace. Proto se hledá způsob, jak odhalit změny chování při co nejmenších změnách atributů. To by umožnilo nejlepší odhad parametrů modelu. Zajímavou možnost představují adaptivní postupy.

Využití adaptivních postupů v průzkumech vyjádřených preferencí (*adaptive stated preference - ASP*) je v poslední době hojně diskutováno. Cílem adaptivního testování je nastavit úroveň atributů alternativ tak, aby co nejvíce stoupal informační přínos následujících pokusů. Informační přínos je dán informační křivkou (*information curve - IC*), která představuje obrácenou hodnotu rozptylu.



Obrázek 13: Logitová a informační křivka

V grafu je znázorněna informační křivka pro logistický model se dvěma alternativami. Hodnota informační křivky pro daný rozdíl užitek obou alternativ je dána násobkem jejich pravděpodobností.

$$IC = p(a1) p(a2) = p(a1) (1 - p(a1)) \quad (10)$$

Z grafu i vzorce vyplývá, že největší informační přínos nastává tehdy, když je rozdíl užitek předkládaných alternativ co nejmenší.

Při využití adaptivních metod tedy z předchozích výsledků odhadneme parametry modelu a pro následující pokus nastavíme hodnoty atributů tak, aby předpokládané užítiky jednotlivých alternativ nabývaly stejných hodnot.

Pomocí adaptivního panelového designu lze získat data pro odhad parametrů jednodušších modelů pro každého respondenta (Richardson, 2002). V tomto designu se kombinuje *conjoint* a standardní *SP* experiment. Takto se adaptivní techniky používají například ve výzkumu preferencí cyklistické infrastruktury cyklisty (Tilahun et al., 2007) nebo v medicíně pro zjišťování preferencí pacientů při léčbě lupenky (Fraenkel et al., 2001) nebo HIV (Beusterien et al., 2005). V dopravě se tento design s výhodou používá při výzkumu rozhodování nákladních přepravců (Fowkes & Shinghal, 2002; Danielis et al., 2005; Fowkes, 2007). O způsobu dopravy nákladu ve firmách rozhodují manažeři logistiky, které je obtížné získat pro dotazování. Navíc různé přepravní úlohy, které tito manažeři řeší, jsou velmi rozdílné, a to především v závislosti na typu nákladu a jeho množství. Proto nedává příliš smysl odhadovat parametry jednoho modelu pro různé přepravní úlohy a manažery různých firem společně, ale je lepší pokusit se parametrizovat dílčí modely a ty poté srovnávat. Tento design byl zkoušen v rámci studie proveditelnosti železničního koridoru Žilina – Košice – Čierna n. T. (Šimeček & Dufek, 2016), kde se ukázal značný rozdíl v přístupu logistiků jednotlivých firem k přepravě různých

komodit. Na rozdíl od obvyklého přístupu, kde respondent volí nejvhodnější alternativu, má zde za úkol podle vhodnosti alternativy seřadit.

**Scenár 1 - úloha 3**

Prepravovaná komodita Potraviný  
Množstvo 10 ton

Odkiaľ Žilina  
Kam Bratislava  
Vzdialenosť 200 km

TRANSPORT RESEARCH CENTRE

Brno Adaptive Social Preference

Zoradte možnosti prepravy tohto nákladu podľa vašej preferencie od najlepšieho (1) po najmenej vhodný (4). Potom pokračujte tlačítkom ĎALEJ.

Cestná preprava		Cestná preprava		Železničná, kombinovaná (kamión - vlak)		Železničná, kombinovaná (kamión - vlak)	
Cena (jednotková)	1.50 EUR/tkm	Cena (jednotková)	1.70 EUR/tkm	Cena (jednotková)	1.20 EUR/tkm	Cena (jednotková)	1.50 EUR/tkm
Čas	12.00 hod	Čas	9.60 hod	Čas	9.60 hod	Čas	12.00 hod
Spoľahlivosť	90.00 % včas a v poriadku	Spoľahlivosť	95.00 % včas a v poriadku	Spoľahlivosť	90.00 % včas a v poriadku	Spoľahlivosť	95.00 % včas a v poriadku
Poradí preferencie	4	Poradí preferencie	3	Poradí preferencie	1	Poradí preferencie	2

Zadanie je v poriadku.

ĎALEJ

Obrázek 14: Návrh experimentu volby přepravního módu v nákladní dopravě na Slovensku (Šimeček & Dufek, 2016)

Adaptivní design lze také použít pro nastavení velikostí atributů v rámci pilotních průzkumů pro klasický *SP* experiment (FOCUS, CDV, 2017), popřípadě pro průběžné „vyлаďování“ velikostí atributů v průběhu *SP* experimentu.

Použití adaptivních technik je nicméně poměrně náročné a zejména v experimentech, které si kladou ambici odhadnout parametry modelu pro každého respondenta, kladou velké nároky na tazatele.

Potřebujete cestovat na vzdálenost asi 50 km mimo město a nemáte k dispozici osobní automobil. Zároveň veřejná doprava pro tuto Vaši cestu rovněž z nějakého důvodu nepřichází do úvahy.

Máte na výběr mezi obyčejným TAXI a čtyřmi variantami autonomního TAXI. Všechny můžete mít přistaveny ve stejný čas, cesta s nimi bude trvat přibližně 60 minut, Vaše pohodlí se nebude rovněž v ničem výrazně lišit.

Podstatné rozdíly mohou být, kromě ceny, pouze v možnosti zasahovat do řízení a v prokázané menší nehodovosti některých autonomních ststémů.

Pro kterou možnost cestování se rozhodnete?

	obyčejné TAXI	autonomní TAXI			
		1	2	3	4
Můžete mít vliv na řízení	-	ANO	ANO	-	-
Nehodovost proti lidskému řidiči	-	MENŠÍ	-	MENŠÍ	-
Cena	1 540 Kč	1 600 Kč	1 400 Kč	1 500 Kč	1 200 Kč
Vaše volba?	A	B	C	D	E

Obrázek 15: Návrh experimentu ochoty cestovat autonomním taxíkem

Jako příklad průběžného nastavování atributů experimentu může sloužit jednoduchý adaptivní design, který navrhujeme pro zjišťování ochoty obyvatel využívat k cestování autonomní taxíky, který představuje nejpravděpodobnější odvětví nástupu autonomních vozidel v osobní dopravě (Brownell & Kornhauser, 2014; Bischoff & Maciejewski, 2016). Respondent má v tomto experimentu za úkol zvolit jednu z pěti alternativ, přičemž alternativou A je normální taxi.

Cena za obyčejný taxík je nastavená podle deklarované vzdálenosti cesty napevno. Ceny za použití autonomních taxíků jsou nejprve stanoveny zkusmo. Po několika pokusech s různými respondenty se odhadnou parametry modelu:

$\beta_a$  - parametr alternativy s autonomním TAXI

$\beta_c$  - parametr kontroly řízení

$\beta_s$  - parametr bezpečnosti

$\beta_{cost}$  - parametr ceny

Následně se přepočtou ceny za jednotlivé alternativy i s autonomním vozidlem (B, C, D, E) podle vzorce:

$$cost_i = \frac{\beta_a + \beta_c control_i + \beta_s safety_i}{\beta_{cost}} + 1540 \quad (11)$$

kde figuruje fixní částka 1540 Kč za cestu obyčejným taxíkem a proměnné  $control_i$  a  $safety_i$  nabývají hodnot nula nebo jedna podle příslušné alternativy. Tímto způsobem je možné pro následující pokusy nastavit atributy ceny pro alternativy s autonomním taxíkem tak, aby si užítky všech alternativ byly pro respondenty co nejvíc podobné, informační přínos dalších odpovědí co největší a aby tak bylo možné co nejlépe odhadovat parametry modelu. Jak bude odpovědí od respondentů přibývat, proces nastavení parametrů ceny se může opakovat.

### 3.5 Výběr respondentů

Způsob výběru respondentů vychází z cílů výzkumu, z charakteru vytvořených scénářů a samozřejmě také z možností, které pro výzkum technicky nebo finančně připadají do úvahy.

Obecně nejlepší pro možnost zobecnění výsledků na populaci je pravděpodobnostní výběr. V takovém případě má každý příslušník cílové populace teoreticky stejnou pravděpodobnost, že se stane respondentem ve výzkumu. Pravděpodobnostní výběr je nicméně finančně nesmírně náročný a v podstatě jej nelze realizovat bez statistické opory, která v nějaké podobě obsahuje údaje o všech příslušnících cílové populace. Stratifikovaný výběr umožňuje navíc počítat s tím, že některých příslušníků populace (strat) se zkoumané rozhodování týká více než jiných. To umožňuje uměle navýšit počet respondentů takto významných skupin nebo naopak snížit počty respondentů ve skupinách, které se rozhodují málokdy (tzv. *over/under-sampling*). Kvótní výběr bývá používán často, protože je relativně



levný a navíc stejně jako stratifikovaný výběr umožňuje cílit na určité skupiny potenciálních respondentů. Nezajišťuje však reprezentativnost výsledků stejně jako příležitostný výběr nebo výběry prováděné metodami snow-ball a podobně.

Ve výzkumech rozhodování se velmi často používá výběr založený na studované volbě (*Choice-Based Sample – CBS*). Respondenti jsou vybíráni přímo z populace osob, které se rozhodují. Tohoto druhu jsou RP data – pozorovaná skutečná rozhodování. Do tohoto typu výběru patří i průzkumy prováděné přímo v dopravě. V rámci studie volby dopravního módu při cestě na letiště (CDV & MEDIAN, 2017) bylo zkoumáno rozhodování leteckých cestujících o způsobu dopravy na letiště Václava Havla v Praze. Cestující byli vybíráni přímo na letišti po té, co cestu na letiště již absolvovali. Mezi respondenty tak byly zastoupeny způsoby dopravy na letiště odpovídajícím způsobem. Pokud by průzkum probíhal například pouze ve veřejné dopravě směřující na letiště, chyběla by ve výběru podstatná část cestujících, a to těch, kteří veřejnou dopravu nezvolili. Průzkumy CBS jsou považovány za velmi efektivní a pro výzkum rozhodování vhodné.

### 3.6 Pilotní studie

Před samotným zahájením průzkumu je vhodné zařadit pilotní studii. Při ní se zjišťují především kvalitativní vlastnosti experimentu, jako jeho realističnost, srozumitelnost a kognitivní náročnost (Johnston et al., 2017). Zároveň lze v průběhu pilotní studie doladit velikosti jednotlivých (především kvantitativních) atributů, aby úrovně mezi sebou rozlišovaly.

Dalším cílem pilotní studie může být ověření procedury dotazování. Ta totiž v případě CBS může poněkud překračovat zaběhnuté postupy výběru respondentů a jejich dotazování. Například v případě studie volby dopravního módu při cestě na letiště (CDV & MEDIAN, 2017) se v rámci pilotního dotazování hledalo místo, kde lze co nejlépe dotazovat letecké cestující (kde letečtí cestující mají čas a čekají a kam mohou získat přístup tazatelé), stejně jako zaměstnance areálu letiště (kde zaměstnanci různých úseků tráví například dobu odpočinku).

## 4. Analýza diskretních voleb

V této kapitole se budeme zabývat odhadem parametrů modelů diskretních voleb. Budeme se také zabývat hodnocením kvality modelů diskretních voleb, což slouží především k volbě takového modelu, který nejlépe odpovídá našim předpokladům o struktuře rozhodovacího procesu a zároveň nejlépe odpovídá datům, na kterých se model odhaduje. V dalších částech kapitoly se budeme zabývat interpretací parametrů modelů, jejich kvantifikací a tím, jak parametry ovlivňují pravděpodobnost volby příslušných alternativ. V interpretacích se zaměříme především na aplikace, které jsou široce používané v dopravě, jako je stanovení hodnoty cestovního času a spolehlivosti dopravy. Z jiných aplikací zmíníme například zjišťování statistické hodnoty života. Na aplikaci analýzy diskretních voleb v parkovací politice si ukážeme, jak změna vlastností alternativ způsobuje změnu poptávky a jak tyto změny v praxi modelovat. Na závěr krátce představíme, jak souvisí modely rozhodování s jednotlivými částmi dopravních modelů a kde se v dopravních modelech využívají výsledky analýzy reálných dat.

### 4.1 Odhad parametrů a hodnocení kvality modelu

Odhad parametrů modelu se provádí metodou největší věrohodnosti. Tato metoda hledá takové nastavení parametrů modelu, aby předpovědi modelu co nejlépe (nejvěrohodněji) odpovídaly skutečným datům naměřeným v experimentu.

Odhad parametrů modelu se provádí pomocí maximalizace funkce věrohodnosti (*likelihood*). Likelihood je součin modelovaných pravděpodobností vybraných ( $w$ ) alternativ pro všechna  $K$  pozorování.

$$L = \prod_{k=1}^K p_{wk} \quad (12)$$

Model, který je dokonale v souladu s daty, bude předpovídat pro alternativy zvolené respondenty pravděpodobnost 1 a pro nezvolené alternativy pravděpodobnost 0. Takto získané pravděpodobnosti jsou agregovány napříč pozorováními tak, že se vzájemně vynásobí. Tím dostaneme vlastně pravděpodobnost toho, že kdyby platil náš model, tak bychom si z náhodné populace vylosovali právě data, která jsme pozorovali. Tato pravděpodobnost je pochopitelně velmi malá, protože si můžeme také vylosovat velké množství různých jiných kombinací výsledků. To se ve vzorci projeví tím, že násobíme mnoho čísel menších než 1. Proto se v praxi pro potřeby odhadu parametrů modelu používá logaritmus věrohodnosti (*log-likelihood*).

$$LL = \ln(L) = \sum_{k=1}^K \ln(p_{wk}) \quad (13)$$

Existuje řada procedur pro odhad parametrů multinomiálních logistických modelů, které se liší v rychlosti a spolehlivosti odhadu. Například program *Biogeme* ve verzi 2.2 obsahuje tři různé metody odhadu, které může uživatel zvolit. Všechny však bývají založené na odhadu parametrů maximalizací věrohodnosti modelu (*Maximum Likelihood Estimation – MLE*), v případě smíšených (mixed) modelů je odhad parametrů řešen pomocí maximalizace simulované věrohodnosti (*Maximum Simulated Likelihood Estimation – MSLE*).

Jedním z ukazatelů kvality modelu je McFaddenovo  $R^2$ :

$$R^2_{McFadden} = 1 - \frac{LL_{model}}{LL_0} \quad (14)$$

kde  $LL_0$  je log-likelihood modelu bez nezávislých proměnných (nulového modelu), který obsahuje pouze alternativně specifické konstanty a  $LL_{model}$  je log-likelihood posuzovaného modelu. McFaddenovo  $R^2$  lze interpretovat podobně, jako standardní koeficient determinace.  $R^2$  říká, jaký díl rozptylu naměřených hodnot model dokáže vysvětlit (Soukup, 2013). McFaddenovo  $R^2$  může nabývat hodnot od mínus nekonečna do jedné. Pokud je záporné, znamená to, že věrohodnost modelu je horší, než věrohodnost nulového modelu. V takovém případě odhad parametrů modelu selhal a je určitě nutné udělat nějaké úpravy. Pokud se hodnota  $R^2$  pohybuje blízko jedné nebo je rovna 1, bývá to obvykle také podezřelé. Podobná situace se může v praxi stát, pokud je odhad prováděn na velmi malém vzorku pozorování, nebo je do modelu zahrnut náhodný parametr, který v rámci odhadu sám „vysvětlí“ veškerý rozptyl. I v takovém případě jde o problém, který je potřeba odstranit. Velikost  $R^2$  se pohybuje obvykle kolem 0,5 až 0,8 pro modely, které opravdu dobře přiléhají na data, ale běžně jsou publikovány také modely, jejichž  $R^2$  je kolem 0,1.

Samotnou hodnotu věrohodnosti nebo jejího logaritmu nelze interpretovat. Závisí nejen na tom, jak model odpovídá datům, na kterých je odhadován, ale také na počtu pozorování a počtu alternativ. Interpretovat lze ale poměr hodnot věrohodností dvou modelů. Test poměru věrohodností (*Likelihood ratio test*) tak slouží k porovnání dvou modelů. Testová statistika má  $\chi^2$  rozdělení se stupni volnosti, které odpovídají rozdílu mezi počtem parametrů modelu 2 a modelu 1.

$$LR = 2 \ln \left( \frac{L_{model2}}{L_{model1}} \right) = 2(LL_{model2} - LL_{model1}) \quad (15)$$

Stejně jako McFaddenovo  $R^2$  neuvažuje test poměru věrohodností počty odhadovaných parametrů modelů. V případě nízkého  $R^2$  lze použít test poměru věrohodností mezi modelem a nulovým modelem k ověření, zda je model (*model2*) i přes malý rozdíl ve věrohodnostech lepší, než nulový model (*model1*). Test poměru věrohodností však slouží nejčastěji k ověření změn v modelu při přidání (vyřazení) proměnné nebo úpravě struktury modelu. Srovnává se tak model po úpravě (*model2*) s modelem před úpravou (*model1*).

Je intuitivně jasné, že se zvyšujícím se počtem odhadovaných parametrů v modelu bude stoupat jeho schopnost přizpůsobit se datům, na kterých je odhadován, a to bez souvislosti

s kvalitou modelu jako takovou. Proto je dobré s počtem parametrů v modelu počítat a modely s více parametry penalizovat. S počtem odhadovaných parametrů pracují informační kritéria Akaikovo ( $AIC$ ) a bayesovské ( $BIC$ ).

$$AIC = 2k - 2LL_{model} \quad (16)$$

$$BIC = \ln(n)k - 2LL_{model} \quad (17)$$





Informační kritéria vyjadřují informační ztrátu, ke které použití modelu vede. Model je tedy tím úspěšnější, čím menší hodnoty v kritériu dosahuje. Obě kritéria započítávají počet odhadovaných parametrů modelu  $k$  a v bayesovském kritériu je navíc zakalkulován počet pozorování  $n$ . Tyto kritéria slouží k porovnání verzí modelů mezi sebou.

## 4.2 Kvantifikace parametrů

Odhadované parametry modelu reprezentují sílu účinku atributu, ke kterému se váží, na pravděpodobnost výběru příslušné alternativy.

Pro některé případy stačí porovnat odhadnuté parametry mezi sebou z pohledu jejich účinku na pravděpodobnost výběru té které alternativy (viz následující kapitola o elasticitě). Tak je tomu například v případě analýzy sebezáchovného chování svědků nehody v přístavu (de Bekker-Grob EW et al., 2015). Zde bylo cílem zjistit především, jak signály (např. dým, čpavý zápach nebo chování ostatních svědků) ovlivňují pravděpodobnost volby různého sebezáchovného chování (útek, nalezení úkrytu nebo zůstání na místě).

Pro jiné případy bude výhodné srovnat odhadované parametry porovnáním přes jeden z nich – tím je obvykle parametr ceny. Představme si model výběru dopravního módu na cestě mezi Prahou a Brnem, odhadovaný na základě  $SP$  experimentu, jehož scénář je ukázaný na obrázku 16.

Jaký z nabízených způsobů dopravy byste využil/a k cestě z Prahy do Brna nebo z Brna do Prahy?			
	<b>Vysokorychlostní železnice</b>		<b>Obyčejná železnice</b>
Cestovní čas	1h 30min	Cestovní čas	2h 28min
Náklady na cestu	400 Kč	Náklady na cestu	263 Kč
	<b>Osobní automobil</b>		<b>Autobus</b>
Cestovní čas	2h 35min	Cestovní čas	2h 50min
Náklady na cestu	326 Kč	Náklady na cestu	165 Kč

Obrázek 16: Scénář  $SP$  experimentu pro volbu způsobu dopravy na trase Praha - Brno

Tento model je zjednodušený, a proto bude mít pro všechny dopravní alternativy stejnou rovnici. Tato rovnice bude pro každou alternativu  $i$ :

$$U_i = ASC_i + \beta_{tt} tt_i + \beta_{cost} cost_i + \varepsilon_i \quad (18)$$

Model má parametr  $\beta_{tt'}$ , asociovaný s atributem cestovního času příslušné alternativy  $tt_i$ . Parametr představuje sílu, s jakou se respondent rozhoduje mezi způsoby dopravy podle toho, jaký cestovní čas mu ta která alternativa dopravy nabízí. Protože při rozhodování vítězí alternativa s nejvyšší hodnotou užitekve funkce, lze znaménko parametru přímo interpretovat. Pokud je znaménko parametru kladné, je příslušný atribut preferovaný a pokud je záporné, je preferované spíše jeho snížení. V případě cestovního času můžeme předpokládat, že čas strávený cestováním představuje spíše ztrátu, a proto parametr u cestovního času budeme očekávat záporný. Stejně tak to bude u parametru  $\beta_{cost'}$  což je parametr atributu ceny  $cost_i$ . Vyšší cena za dopravu sama o sobě představuje vyšší náklady a tedy nižší prospěch z příslušné alternativy.

Užitek  $U$  z každé alternativy je bezrozměrné číslo, a proto parametry v užítkových rovnicích mají obrácené jednotky ke svým atributům. Takže například, jestliže cestovní čas je v minutách, má parametr  $\beta_{tt}$  jednotku [1/min]. Díky tomu lze použít jednotku, jejíž parametr se podařilo robustně odhadnout, ke kvantifikaci účinku atributů alternativ převedením na tuto jednotku. Nejčastěji se jako referenční jednotka používá cena. Ostatně modely volby jsou mikroekonomickými modely. V takovém případě se jedná o převod velikosti účinku atributů na peníze – tedy monetarizaci. Může jít ale v principu o jakoukoli jinou jednotku. Hovoří se pak o takzvané ochotě platit (*WTA - willingness to pay*) za příslušný atribut, například službu, nebo naopak ochotě akceptovat (*WTC - willingness to accept*) určité zhoršení za finanční náhradu (Henscher et al., 2005; Shogren, 1994; Hanemann, 1991). Ochota platit a ochota akceptovat nemusí být totéž s obráceným znaménkem, a to díky teorii prospektu (viz kapitola 2.1), podle které lidé, kteří se rozhodují za nejistoty, jinak vnímají prostředky vynaložené za volbu lepší alternativy a jinak prostředky ušetřené za nějakou nevýhodu. To se projevuje i v SP experimentech, kde rovněž, v souladu s teorií prospektu, respondenti většinou akceptaci nějaké ztráty subjektivně hodnotí jako závažnější než naopak přesně stejný přínos (Hess et al., 2008).

Podíl parametru času ( $\beta_{tt}$ ) a parametru ceny ( $\beta_{cost}$ ) se nazývá hodnota času (*Value of Time – VoT*):

$$VoT = \frac{\beta_{tt}}{\beta_{cost}} \quad (19)$$

Hodnota času je jedním z klíčových ukazatelů volby dopravního módu a je to také velmi významný vstup do ekonomického posuzování přínosů dopravních staveb. Dejme tomu, že cestovní čas je v minutách a cena za dopravu v korunách. Potom hodnota času znamená, nakolik si cestující cení v korunách jednu minutu svého času stráveného cestováním. Hodnoty cestovního času jsou většinou uváděny v příslušné měně za hodinu. Proto, pokud by cestovní čas byl v minutách (tak bývá často vyjádřen v průzkumech), je potřeba hodnotu času násobit počtem minut v hodině.

$$VoT = 60 \frac{\beta_{tt}}{\beta_{cost}} \quad (20)$$

Jednotlivé parametry modelu jsou odhadovány s určitou chybou – chybou odhadu. Tato chyba bývá vyjádřena jako standardní chyba odhadu (*Standard Error - SE*). Rovněž pro kvantifikované veličiny lze dopočítat standardní chybu, a to ze standardní chyby odhadu příslušného parametru, standardní chyby parametru použitého ke kvantifikaci a kovariance mezi těmito parametry. Tyto údaje jsou ve výstupech modelu obvykle k dispozici. Výpočet chyby vychází z obecného vzorce určení výsledné chyby pro nepřímá měření (Tichá, 2004).

Pro hodnotu času (viz výše) pak přibližně platí, že jeho standardní chybu odhadu lze vypočítat z podílu standardních chyb odhadu parametrů času ( $SE(\beta_{tt})$ ) a ceny ( $SE(\beta_{cost})$ ), a dále kovariance těchto parametrů  $cov(\beta_{cost}, \beta_{tt})$  podle vzorce:

$$SE_{VoT} \cong \sqrt{\left(\frac{1}{\beta_{cost}} SE_{\beta_{tt}}\right)^2 + \left(\frac{\beta_{tt}}{\beta_{cost}^2} SE_{\beta_{cost}}\right)^2 - 2 \frac{\beta_{tt}}{\beta_{cost}^3} cov(\beta_{cost}, \beta_{tt})} \quad (21)$$

Vzorec platí pouze pro chyby, které se limitně blíží nule. Pro relativně malé chyby jsou však rozdíly vypočítané chyby proti skutečnosti zanedbatelné. To se dá říci, pokud se chyby odhadu obou parametrů pohybují pod hranicí statistické významnosti  $p < 0,05$ .

Započtení kovariance do odhadu chyby kompozitní proměnné může vést ke zmenšení nebo ke zvětšení odhadované chyby. Pokud mají oba parametry souhlasné znaménko, pak kladná kovariance odhadnutou chybu kompozitní proměnné snižuje a naopak záporná kovariance odhadnutou chybu zvyšuje. Pokud parametry souhlasné znaménko nemají, bude to naopak.

V prostředí *R* může funkce pro výpočet kvantifikovaných parametrů a jejich standardních chyb vypadat například následovně:

```

quantif=function(model,refname) {
  km=data.frame(val=coef(model),
                se=sqrt(diag(vcov(model,"coefficient",type=c(,"cov")))),
                cov=vcov(model,"coefficient",type=c(,"cov"))[refname,]),
  refnum=which(names(coef(model))==refname)
  km$qval=km$val/km$val[refnum]
  km$qse=sqrt(
    (1/km$val[refnum]*km$se)^2
    +(km$qval/km$val[refnum]*km$se[refnum])^2
    -2*km$qval/km$val[refnum]^2*km$cov
  )
  return(km)
}

```

Funkce vrací tabulku odhadnutých parametrů, jejich chyby, kovariance, kvantifikaci parametrů a vypočtenou chybu této kvantifikace z modelu *model* a podle nezávislé proměnné (atributu) *refname*. Tou je většinou proměnná označující cenu, ale nemusí to tak být vždy.

Hodnota času je jedna z komponent takzvaných generalizovaných cestovních nákladů. Generalizované cestovní náklady jsou sumou jak finančních, tak i nefinančních nákladů na cestu. Typickým nefinančním nákladem na cestu je právě cestovní čas, jehož finanční

hodnotu můžeme získat pomocí výše uvedené kvantifikace. V DeSerpově teorii je cestovní čas nahlížen jako komodita. Naopak, v Beckerově teorii je cestovní čas vnímán jako zdroj k získání produktu, kterým je výsledek cestování – přemístění se. To má důsledky pro rozdílný způsob odhadu ceny cestovního času. V rámci DeSerpovy teorie by cena cestovního času byla ve všech případech stejná (*total utility*) a odhadovala by se tak stejná pro všechny případy cest. Z Beckerovy teorie naopak vyplývá, že hodnota času je závislá na konkrétním případě cesty (například majetku cestujícího nebo potřebnosti příslušné cesty, která souvisí s jejím účelem) a proto by měla být hodnota času odhadována zvlášť pro tyto různé případy jako tzv. *marginal utility* (Jara-Díaz, 1998). Truong a Hensher (1985) se přiklání k Beckerově teorii. Dokazují, že se hodnota cestovního času liší podle konkrétních okolností cestování. Dnes se hodnota cestovního času obvykle odhaduje zvlášť pro různé účely cest a také pro cestování různými dopravními prostředky.

Hodnota cestovního času se používá například ve studiích proveditelnosti pro vyčíslení přínosů dopravních staveb. Pokud realizace dopravní stavby povede ke zrychlení dopravy, dojde i ke zkrácení cestovního času (Hensher, 2001; De Jong et al., 2007; Hess et al., 2008). V této souvislosti se hovoří o hodnotě uspořené času (*VTTS – Value of Travel Time Savings*), což je právě čas, který uživatelé dopravní infrastruktury uspoří zkrácením cestovního času a který proto mohou alokovat na jiné účely. Hodnota uspořené času tvoří podstatnou část zjišťovaných přínosů budované dopravní infrastruktury. Ve vyspělých zemích je to až 80 % z celkového benefitu (Athira et al., 2016). Celkové úspory (dejme tomu za rok) se pomocí VoT získají jednoduše tak, že se touto hodnotou násobí průměrná úspora času, která vznikne realizací stavby na jednu cestu a počtem takových cest za rok. Metz (2008) namítá proti přílišnému důrazu na úsporu času u studií proveditelnosti, že zrychlení cestování reálně k časové úspoře nevede. Lidé mají podle něj tendenci trávit cestováním stále stejné množství času. Místo toho vede zrychlení cestování na nové dopravní infrastrukturu k přínosům z možnosti dosažení vzdálenějších destinací. Metz proto navrhuje měřit místo hodnoty cestovního času hodnotu dostupnosti (*value of access*).

Dalším významným ukazatelem přínosu dopravních staveb je hodnota spolehlivosti dopravy (*VoR – Value of Reliability*). Je to cena, kterou jsou ochotni cestující zaplatit za zvýšení spolehlivosti dopravy. Spolehlivost dopravy lze definovat různě. Spolehlivost cestovního času lze chápat (a bývá tak vyjadřována) jako směrodatnou odchylku variability skutečného cestovního času (de Jong et al., 2004). To je ovšem koncept nepoužitelný pro SP průzkumy, protože v rámci něho nedokáže většina lidí uvažovat. Proto se v dotaznících spolehlivost cestovního času předkládá v jako výpis možných cestovních časů (Hensher, Li & Rose, 2013) nebo jejich zobrazení do podoby histogramů (Tilahun & Levinson, 2010), nebo jako konfidenční intervaly, v rámci kterých se cestovní časy vyskytují s určitou pravděpodobností, tedy jako šance, že skutečná doba cesty bude odpovídat předpokládané době s menším spožděním než například 15 minut (de Jong et al., 2007; Small, 2012; Carrion & Levinson, 2012; Hensher, Li & Rose, 2013). Z těchto ukazatelů je pak směrodatná odchylka cestovního času odvozena. Hodnota spolehlivosti cestovního času bývá často vyjadřována vzhledem



k hodnotě cestovního času (Carrion & Levinson, 2012). V takovém případě se spolehlivost vyjadřuje jako poměr mezi hodnotou spolehlivosti a hodnotou cestovního času (*RR-reliability ratio*). Hodnota spolehlivosti nabývá na důležitosti zejména při modernizacích stávající dopravní infrastruktury. Po modernizaci (například po přidání jízdních pruhů na dálnici) většinou nelze očekávat podstatné zkrácení cestovního času a z toho pramenící úspory cestovního času. Místo toho můžeme očekávat úbytek dopravních kongescí a s tím spojené zvýšení časové spolehlivosti dopravy. Hodnota času stráveného v dopravní kongesci nebo při zpoždění vlaku je vnímána podobně jako samotná hodnota času (de Jong et al., 2004; Carrion & Levinson, 2012) a začíná být zkoumána i u nás (Máca et al., 2012). Riziko zpoždění toriž vede cestující k tomu, aby si vytvářeli časovou rezervu (de Jong et al., 2007).

V nákladní dopravě je hodnota spolehlivosti rozdílná v závislosti na přepravovaném nákladu, ale v zásadě je výrazně vyšší než hodnota času přepravy (de Jong et al., 2009; Shams, 2016). Ve studii proveditelnosti modernizace železničního koridoru Žilina – Košice – Čierna n. T (Star EU & CDV, 2014) se samotný cestovní čas v nákladní dopravě ukázal téměř nedůležitý proti například ceně za přepravu nákladu. Cestovní čas hrál podstatnější roli pouze u komodity potravin. Zato značnou roli hrála spolehlivost přepravy, tedy šance, že náklad dorazí na místo určení v pořádku a včas. Ukazatel spolehlivosti byl naprosto zásadní zejména u kontejnerů, protože kontejnerový způsob přepravy nákladu často směřuje do přístavů na nákladní lodě. V takovém případě by zpoždění zásilky mohlo znamenat, že nestihne být v přístavu odbavena na loď, což by pro přepravce znamenalo ohromné ztráty. Také situace, ve které takový náklad do přístavu dorazí předčasně, generuje ztráty přepravci, spojené se zařizováním a platbou krátkodobého uskladnění nákladu. Vysoká spolehlivost a také v nákladní dopravě je také naprosto nezbytnou podmínkou pro režim výroby *just-in-time* (Fowkes et al., 2004; Tseng, Yue & Taylor, 2005), ve kterém se pomocí správného načasování přísunu výrobních vstupů a odběru výstupů z výroby snižují náklady na skladování. Z tohoto důvodu mnozí nákladní přepravci rezignují na případné výhody železniční dopravy, která jim připadá pro režim *just-in-time* málo flexibilní a spolehlivá (Star EU & CDV, 2014).

Rychlost přepravy, tedy hodnota cestovního času, i její spolehlivost hrají roli také ve veřejné dopravě, kde svoji roli sehrávají také další podstatné vlastnosti veřejných přepravních služeb. Veřejní dopravci (například městští, železniční nebo autobusoví) se snaží koncipovat své služby tak, aby uspokojili a přitáhli co nejvíce potencionálních zákazníků. Další zkvalitňování hlavních přepravních služeb jako je rychlost a spolehlivost není možné do nekonečna a stojí provozovatele služby stále více prostředků. Také každé zlepšení stávajících služeb nebo zavedení nové služby navíc bude muset nakonec zaplatit zákazník. Analýza diskrétních voleb umožňuje finančně „nacenit“ hodnoty těchto služeb tak, aby odpovídaly hodnotám, které služby mají pro jejich uživatele. Dopravce se tak díky průzkumu dozví, kolik je pasažér ochotný zaplatit za tu kterou službu nebo její zlepšení. Průzkum tohoto typu je většinou potřeba udělat formou vyjádřených preferencí, kdy respondent vybírá mezi hypotetickými alternativami dopravy, kam mohou být zakomponovány i dosud nerealizované služby a jejich nové vlastnosti.



Analýza ochoty platit za služby tak může veřejnému dopravci odpovědět na otázky typu:

*Kolik si jsou zákazníci ochotní připlatit navíc za zkrácení cestovního času průměrně o 10 minut?*

*Kolik si jsou uživatelé městské hromadné dopravy ochotni zaplatit za jistotu, že si ve vozidle sednou?*

*Kolik si jsou zákazníci ochotní zaplatit navíc, pokud se frekvence spojů veřejné dopravy zvýší průměrně o 10%?*

*Kolik si jsou zákazníci ochotní zaplatit navíc, pokud by se průměrný počet přestupů ve veřejné dopravě průměrně snížil o 0,1?*

Významným faktorem pro volbu dopravního módu je komfort v širokém slova smyslu. Ve vozidlech veřejné dopravy je daný především obsazeností vozidla a také jeho výbavou. V rámci studie modernizace železničního koridoru Žilina – Košice – Čierna n. T. (Star EU & CDV, 2014) jsme například zjistili, že za jistotu nalezení místa k sezení ve veřejné dopravě, jsou lidé ochotni zaplatit 1,9 EUR navíc v autobuse a 1,77 EUR ve vlaku, a to na cestách od 20 do 100 km. Jsou také ochotni zaplatit 0,99 EUR navíc za cestování v moderním voze s WiFi připojením. Ochota platit za WiFi připojení patrně závisí na celkovém pokrytí veřejnou sítí a také na délce cesty. Pokud je pokrytí značné anebo délka cesty relativně krátká, ochota platit za WiFi se neobjeví. To se ukázalo ve studii modernizace železničního spojení Praha – letiště Praha – Kladno (CDV & MEDIAN, 2017), kde tuto službu považovali respondenti za málo užitečnou nebo za samozřejmost.

Ve studiích, které zahrnují veřejnou dopravu opakovaně zjišťujeme značný význam přestupů ve veřejné dopravě (CDV & FOCUS, 2017; CDV & MEDIAN, 2017). Respondenti z Brna a brněnského zázemí jsou ochotni zaplatit  $6,1 \pm 2,7^2$  Kč za snížení počtu přestupů o jeden, přičemž u obyvatel Brna se taková ochota nevyskytuje, ale obyvatelé brněnského zázemí jsou ochotni za snížení počtu přestupů o jeden platit až  $21,7 \pm 11,6$  Kč. To je patrně způsobeno dobrou návazností a frekvencí spojů v Brně městě a naopak riziky, které při přestupu ve veřejné dopravě mohou nastat v rámci cestování brněnským zázemím. Tyto potíže, spojené s přestupy může značně snížit kvalitní koordinace veřejné dopravy, což se také v Jihomoravském kraji daří (viz také kapitola 7). V případě cestování na letiště (CDV & MEDIAN, 2017) hraje počet přestupů také zřejmě zásadní roli, i když jeho význam bylo obtížné určit. Pohyboval se od desítek po stovky korun za přestup. Je to proto, že v případě cestování na letiště jsou rizika spojená s cestou a tedy i přestupy vážená rizikem zmeškání letu. U obyvatel Kladenska na cesty do Prahy byla ochota platit za snížení počtu přestupů o jeden  $17,5 \pm 4,7$  Kč.

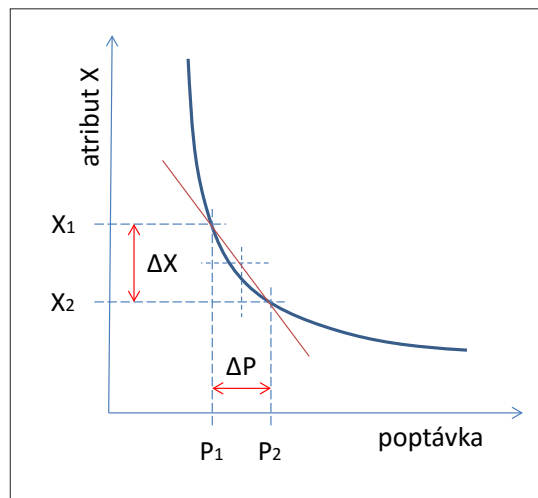
Kromě veřejných dopravců mohou využít i jiní činitelé v dopravě průzkumy založené na vyjádřených preferencích a analýze diskrétních voleb. Tyto průzkumy mohou sloužit například k formulaci parkovacích politik, jako je nastavování ceny parkování P+R (park and

ride). Zatímco parkovací politika v centrech měst je spíše čistě politickou záležitostí, P+R je služba, kterou jednoznačně potřebujeme nastavit tak, aby byla využívána co nejvíce. Proto potřebujeme nastavit cenovou hladinu parkování tak, aby byla pro uživatele přijatelná. Cílem je naplánovat takovou kapacitu P+R parkovišť a zároveň definovat takovou cenovou politiku, aby se kapacity P+R parkování naplnily. K tomu slouží zavedení parametrů pro cenu parkování do volby dopravního módu (resp. řetězce dopravních módů) dopravního modelu. Tím parametrem může být právě ochota platit za P+R parkování. Také může být vhodné analyzovat kromě ochoty platit za parkování elasticitu motoristů na cenu za parkování (viz následující kapitola).

Ochota platit respektive akceptovat se může týkat také bezpečnosti. Pomocí průzkumů vyjádřených preferencí se hodnotí vliv bezpečnosti na volbu cestovní trasy (de Blaeij et al., 2003; Rizzi & de Dios Ortúzar, 2003; Iragüen & de Dios Ortúzar, 2004) nebo ochota platit za opatření, která dopravní bezpečnost zvyšují (Jamson et al., 2008). Pomocí scénářů s dopravně-bezpečnostními opatřeními nebo jinými opatřeními, které zvyšují bezpečnost (Mrozek & Taylor, 2002; Viscusi & Aldy, 2003) se naceňuje takzvaná statistická hodnota života (VSL – *Value of Statistical Life*), používaná například v pojišťovnictví. Jde o obecnou hodnotu lidského života bez ohledu na věk, postavení a další charakteristiky. Tato hodnota byla několikrát stanovována také pro Českou republiku a například ve studii Alberini a Ščasný (2011) vychází na 18,2 mil. Kč.

### 4.3 Elasticita

Elasticita v modelu vyjadřuje vliv určitého atributu na poptávku, tedy na modelovanou pravděpodobnost výběru určité alternativy. Jde o změnu poptávky po příslušné alternativě při změně daného atributu alternativy. Změna poptávky po alternativě a změna atributu je vyjádřena v procentech.



Obrázek 17: Elasticita - vztah mezi změnou atributu a změnou poptávky

V situaci, kdy je atribut  $X$  na úrovni  $X_1$ , má poptávka (podíl alternativy vůči ostatním alternativám) velikost  $P_1$ . Při změně atributu  $X$  o  $\Delta X$  (na hodnotu  $X_2$ ) se změní poptávka o  $\Delta P$  (tedy na hodnotu  $P_2$ ).

Rozlišujeme dva typy elasticity. Bodovou a úhlovou elasticitu.

Bodová elasticita má výchozí bod v současném stavu ( $P_1$  a  $X_1$ ) a její rovnice je:

$$E_{point} = \frac{\Delta P}{\Delta X} \frac{X_1}{P_1} \quad (22)$$

Bodová elasticita je vhodná tehdy, když hledáme rozdíly od počátečního stavu. Ptáme se například, jak se změní využívání osobního automobilu oproti současnému stavu, pokud zvýšíme cenu za parkování ze současných 30 Kč na 40 Kč (o 10 Kč).

Úhlová elasticita má výchozí bod uprostřed intervalu  $X_1, X_2$  a  $P_1, P_2$ . Její vzorec je:

$$E_{arc} = \frac{\Delta P}{\Delta X} \frac{X_1 + X_2}{P_1 + P_2} \quad (23)$$

Úhlová elasticita je obecnější. Můžeme se ptát, jak se změní využívání automobilu, pokud zvýšíme cenu za parkování. Odpověď pak není tolik závislá na počáteční úrovni atributu a poptávky.

Dosud jsme se ptali, jak změna nějakého atributu ovlivní poptávku po alternativě, které se tento atribut týká. To je přímá (*direct*) elasticita. Také se můžeme ptát, jak změna atributu nějaké alternativy ovlivní poptávku po jiných alternativách. Tomuto typu elasticity se říká křížová (*cross*) elasticita. Tímto způsobem se dá například odhadnout, jak změna ceny za veřejnou dopravu ve městě může ovlivnit využívání individuální automobilové dopravy.

Při interpretaci odhadnutých parametrů modelu se vypočítá *direct-point* elasticita atributu  $i$  alternativy  $a$  podle vzorce:

$$E_{ai} = \beta_{ai} x_{ai} (1 - P_a) \quad (24)$$

kde  $\beta_{ai}$  je parametr atributu  $i$  alternativy  $a$ ,

$x_{ai}$  je současná velikost atributu  $i$  alternativy  $a$ ,

$P_a$  je podíl alternativy  $a$  mezi všemi alternativami (poptávka).

Cross-point elasticitu atributu  $i$  alternativy  $a$  pak dostaneme ze vzorce:

$$E_{j \neq a, i} = -\beta_{ai} x_{ai} P_a \quad (25)$$

kde  $\beta_{ai}$  je parametr atributu  $i$  alternativy  $a$ ,

$x_{ai}$  je současná velikost atributu  $i$  alternativy  $a$ ,

$P_a$  je podíl alternativy  $a$  mezi všemi alternativami (poptávka).

Když je parametr atributu kladné číslo, elasticita je rovněž kladná a se zvyšující velikostí atributu se poptávka po příslušné alternativě zvyšuje. Poptávka po ostatních alternativách se naopak snižuje. U záporné velikosti parametru je to naopak.

Abychom mohli z odhadnutých parametrů modelu počítat elasticity, musíme znát současnou poptávku po jednotlivých alternativách, což se obvykle musí nějak zjistit. Například ke zjišťování podílu dopravních módů se používají standardizované cestovní deníky.

V praxi se často používá inkrementální multinomiální model (Ben-Akiva & Lerman, 1985; Martin & McGuckin, 1998), který současně počítá nové velikosti poptávek pro všechny alternativy. Inkrementální model je formulovaný takto:

$$P'_a = \frac{P_a \times e^{\Delta u_a}}{\sum_{i=1}^k (P_i \times e^{\Delta u_i})} \quad (26)$$

kde  $P'_a$  je počítaná poptávka po alternativě  $a$ ,

$P_a$  je původní poptávka po alternativě  $a$ ,

$\Delta u_a$  je změna užítka alternativy  $a$ ,

$P_i$  je původní poptávka po alternativě  $i$ ,

$\Delta u_i$  je změna užítka alternativy  $i$ .

Změna užítka alternativy  $a$  se vypočítá jednoduše jako násobek parametru a změny příslušného atributu.

$$\Delta u_a = \beta_{an} \Delta x_{an} \quad (27)$$

V prostředí R lze použít k výpočtu jednoduchého inkrementálního modelu následující funkci:

```
incremental = function(share, alt, beta, dx) {
  inc = share
  inc[alt]=inc[alt]*exp(beta*dx)
  inc=inc/sum(inc)
  return (inc)
}
```

Tato funkce napsaná pro R má parametr *share*, což je vektor současných velikostí poptávky (podílů na trhu) jednotlivých alternativ. Parametr *alt* určuje, ve které alternativě dochází ke změně užítka. Parametry *dx* a *beta* jsou změna velikosti atributu a jeho odhadnutý parametr v modelu. Výstupem je vektor nových poptávek.

Jedná se o nejjednodušší verzi inkrementálního modelu, kde se mění pouze jeden atribut u jedné alternativy. Snadno lze napsat funkci, která bude počítat nové poptávky pro kombinace změn atributů v různých alternativách. Výše uvedený inkrementální model však bude pro potřeby této knihy stačit.

Ve studii proveditelnosti modernizace železničního koridoru Žilina – Košice – Čierna n.T. (Star EU & CDV, 2014) bylo pro realizátory modernizace důležité zrychlení železniční dopravy. Tvůrci studie se domnívali, že právě toto zrychlení přivede na železnici zákazníky a přinese tak největší benefity. Aby bylo zrychlení železniční dopravy na tomto koridoru proveditelné, vyžadovalo by podstatné napřimování jednotlivých úseků tratě ve složitém terénu. Toto napřimování by si vyžádalo značné investiční náklady na zbudování mostů a ražbu tunelů. Investoři se nicméně domnívali, že se to vyplatí. Byla provedena analýza, jak by zákazníci železniční dopravy reagovali na změnu cestovního času. Ukázalo se, že na krátkých cestách (20 km) by se cestovní čas musel průměrně zkrátit o 15 minut, aby se podíl železniční dopravy zvýšil o 10% proti současnému stavu (z 19,2% na 29,2%). Pro dlouhé cesty (100 km) by se podíl železniční dopravy zvýšil o 10% (ze současných 20% na 30%), kdyby se cestovní čas zkrátit průměrně o 45 minut. Tyto výsledky korespondují s nízkou hodnotou času ve vlaku a tedy malou citlivostí potenciálních železničních pasažérů na změnu cestovního času. I tento výsledek nakonec vedl investory ke zvolení skromnější varianty modernizace a přinutil je klást větší důraz na zvyšování spolehlivosti železniční dopravy.

Ve studii proveditelnosti pro modernizaci železnice Praha – letiště Praha – Kladno (CDV & MEDIAN, 2017) se počítalo také se zvážením přínosu zavedení P+R parkování a zjištění účinku zavedení rezidenčního parkování v Praze. K tomu měl sloužit především dopravní model, do kterého vstupoval parametr vlivu ceny za parkování na volbu způsobu dopravy. Prováděla se ale také analýza elasticity.

Průzkum cestujících z Kladenska do Prahy dovoval odlišit tři možnosti parkování na cestě do Prahy, a to parkování mimo území Prahy (respondent dál pokračuje v cestě do Prahy meziměstskou veřejnou dopravou, na okraji Prahy (respondent dál pokračuje pražskou městskou dopravou) a parkování v Praze. Ceny za parkování v těchto třech možnostech hrají roli při rozhodování o způsobech dopravy a tyto způsoby dopravy společně ovlivňují. Parkování v Praze končilo 22 % cest z průzkumu, na okraji Prahy to byly 4 % a mimo Prahu 3 %, ostatní automobil nepoužili. Z inkrementálního modelu vyplynulo, že zvýšení ceny za parkování v centru Prahy (zavedení rezidenčního parkování) má vliv především na volbu dopravního módu směrem k veřejné dopravě a jen malý vliv na to, zda cestující využijí P+R parkování na okraji Prahy. Je to proto, že v době průzkumu byly možnosti parkování mimo Prahu a pokračování v cestě veřejnou dopravou prakticky nedostupné nebo neatraktivní (4 a 3 % uvedená výše). Aby tito cestující začali využívat P+R parkování, potřebují, aby tato možnost společně s napojením na veřejnou dopravu začala reálně existovat.

U zaměstnanců v areálu letiště Praha, kteří se do zaměstnání dopravovali autem v 30 % a parkovali v areálu, se například ukázalo, že by nárůst průměrné ceny za parkování o 2000 Kč za rok znamenal snížení cest do práce autem z 30 % na 22 % ve prospěch cestování veřejnou dopravou. Nejvíce jsou na tuto změnu citlivé ženy. Nárůst průměrné ceny za parkování o 2000 Kč za rok bude u zaměstnankyň areálu letiště znamenat pokles cest do práce autem z 26 % na 16 %. Rovněž velmi citliví na cenu parkování byli zaměstnanci do 30 let, ti se ale již

nyní na letiště dopravují automobilem jen málo, a proto účinek zvýšení nákladů za parkování u nich nemá takový absolutní efekt.

V zahraničí je využívání průzkumů vyjádřených preferencí pro nastavení parkovacích politik široce užívané již od 90. let (Axhausen & Polak, 1991; Widmer, 2004; Dell'Olio, 2009).

#### 4.4 Interpretace koeficientů u smíšených modelů

Představme si, že v modelu volby dopravního módu budeme předpokládat, že parametr času je náhodná proměnná. Odpovídalo by to předpokladu, že různí lidé kladou různý důraz na dobu strávenou na cestě a podle toho si vybírají způsob cestování. Proto budeme pro parametr času odhadovat distribuci ve smíšeném modelu. Pokud odhadovanou distribuci nebudeme omezovat (budeme například předpokládat normální rozdělení) znamená to, že parametr času může individuálně nabývat také kladných hodnot. To by znamenalo, že kdybychom pro takové případy počítali hodnotu času, bude záporná. To je v rozporu s intuicí o racionálním chování. Hess a kolektiv (2005) argumentují zjištěními v různých studiích, že za určitých okolností hodnota času skutečně záporná být může. Cestování může být samo o sobě aktivitou a potřebou lidí, negativní užitek z trávení času cestováním může být také kompenzován přínosem nějaké aktivity vykonávané současně. Mnoho cestujících dává přednost časově delší cestě po železnici před autem nebo letadlem, protože čas ve vlaku mohou využít k práci nebo naopak k relaxaci. Za negativní hodnotou času může být také nepozorovaná preference, týkající se například komfortu, potěšení z řízení automobilu a podobně (Hess et al., 2005).

Dalším problémem, které smíšené modely přináší je poněkud umělé zvýšení elasticity. Použití smíšeného modelu nemá vliv na podíly jednotlivých koeficientů, ale ovlivňuje jejich absolutní velikosti. Je to způsobeno tím, že část chyby modelu vysvětlí náhodná proměnná (resp. proměnné) v modelu, které ošetřují individuální rozdíly v preferencích, a právě rozdíly preferencí mezi individui snižují celkovou elasticitu. Pak změna parametru v těchto modelech může vést k poměrně velké změně pravděpodobností volby alternativ. S tímto problémem jsme se setkali při tvorbě modelu volby dopravního módu na cestách z Kladenska do Prahy, kde respondenti odpovídali na čtyři rozhodovací úlohy, a proto musel být odhadnut smíšený model pro panelová data (CDV & FOCUS, 2017). Pro potřeby modelu volby dopravního módu pro dopravní model je možné náhodné proměnné v modelu přímo simulovat, pokud to modelovací software dovoluje, nebo elasticitu opětovně snížit tím, že se všechny koeficienty modelu úměrně sníží.

## 5. Případová studie: Výběr pronájmu studentského bydlení

V Brně je velmi rozšířený trh s pronájmy studentského bydlení, protože je zde velké množství studentů a kapacity studentských kolejí jsou naprosto nedostatečné. Charakter nabízených pronájmů je různý a také způsob jejich popisu v inzerátech se liší. S frekventanty kurzu analýzy diskrétních voleb na FSS MU jsme se rozhodli prozkoumat preference studentů, kteří si hledají pronájem. Kromě ceny se měl experiment zaměřit také na další aspekty bydlení, které jsou pro studenty důležité.

### 5.1 Příprava experimentu

Před započítím přípravy experimentu proběhl rozhovor mezi samotnými výzkumníky, kterými byli studenti psychologie na FSS MU. Tito studenti by mohli patřit mezi osoby, které studentský pronájem skutečně hledají. Cílem rozhovoru bylo nalézt relevantní aspekty volby pronájmu studentského bydlení, které by stálo za to do experimentu zahrnout. Tento rozhovor nahrazuje *focus-group*, která je vhodná pro nalezení vhodných atributů a nastavení jejich úrovní.

Skupina vybrala tyto atributy a k nim příslušící úrovně:

- Měsíční cena pronájmu na osobu – 2200-2700 Kč, 4800-5300 Kč a 7400-7200 Kč
- Počet osob na pokoji – 1, 2 a 3 osoby na pokoji
- Zařízení pokoje – nezařízený, zařízený, nově zařízený
- Vzdálenost od školy – 10 minut, 30 minut

V rámci každé úrovně ceny pronájmu hodnota náhodně variuje v rozsahu 500 Kč, aby se v jednom scénáři nevyskytovaly zcela stejné ceny, což by vypadalo zvláště.

Skupina dále rozhodla, že se jeden scénář bude skládat ze čtyř alternativ a každému respondentovi se budou nabízet tři různé scénáře. Jako nástroj pro webovou administraci dotazníku byl zvolen software *Qualtrics*<sup>3</sup>.

Počet osob na pokoji: 2 Zařízení bytu: nově zařízený Vzdálenost od školy: 30 minut Cena: 7500Kč	Počet osob na pokoji: 1 Zařízení bytu: nezařízený Vzdálenost od školy: 10 minut Cena: 5300Kč	Počet osob na pokoji: 2 Zařízení bytu: zazařízený Vzdálenost od školy: 10 minut Cena: 2400Kč	Počet osob na pokoji: 3 Zařízení bytu: nezařízený Vzdálenost od školy: 10 minut Cena: 2200Kč
--	---	---	---

Obrázek 18: Design experimentu

Z počtu zamýšlených atributů a jejich úrovní vyplývá, že všech kombinací je  $3 \times 3 \times 3 \times 2 = 54$ . Protože těchto kombinací je málo a protože nebylo dopředu známo, kolik respondentů se podaří získat, rozhodli se výzkumníci, že použijí plně faktoriální design.

## 5.2 Výběrový soubor

Celkem bylo získáno 234 pozorování (jednotlivých rozhodnutí) od 78 respondentů. Respondenti se rekrutovali nahodilým výběrem z facebookových skupin v období duben až květen 2018. Základní sociodemografické údaje výběrového souboru jsou uvedeny v tabulce.

Tabulka 5: Sociodemografie respondentů

	n	%
Celkem	78	100
Pohlaví		
muži	26	33
ženy	52	67
Dosažené vzdělání		
základní	0	0
střední odborné	0	0
střední s maturitou	33	42
bakalářské	35	45
magisterské/inženýrské	5	6
doktorské	5	6
Zaměstnání (alespoň 0.2 úvazku)		
ano	72	92
ne	6	8

Ambicí výzkumníků bylo zjišťovat také věk respondentů, to se ale z technických důvodů nepodařilo.

## 5.3 Příprava dat

Skript pro *R* a tabulka dat jsou v souborech „model\_privat.R“ a „data\_privat.dat“ (<https://www.cdv.cz/vyzkum-diskretnich-voleb/>). Data jsou ve formátu vhodném pro analýzu v programu *Biogeme*. Jde o formát tzv. „wide“, kde na každém řádku je jedno pozorování. Je zde číslo respondenta (*id*), údaje o respondentově pohlaví, vzdělání a zaměstnání (*gen*, *edu*, *zam*), číslo pokusu (*trial*). Dále následují proměnné jednotlivých alternativ, což je počet osob na pokoji (*posob*), zařízení pokoje (*zariz*), vzdálenost privátu od školy (*vzdal*) a výše měsíčního nájmu na osobu (*najem*). Proměnné pro jednotlivé alternativy jsou označeny číslem alternativy za podtržítkem. Po té následuje sloupec s číslem vybrané alternativy (*Choice*). Nakonec je proměnná (*ok*), která označuje řádky dostatečně kompletní pro další analýzu.

Pro analýzu v prostředí *R* je potřeba data upravit. Úprava původní tabulky probíhá ve dvou krocích. Tabulka je načtena a odstraněny jsou nekompletní řádky (podle proměnné *ok*). Proměnné *zariz* a *posob* jsou nominální proměnné, a proto jsou převedeny na faktory.



```
d = read.table(„data_privat.dat“, header = T)
d = d[d$ok==1,]
d$zariz_1 = as.factor(d$zariz_1)
d$zariz_2 = as.factor(d$zariz_2)
d$zariz_3 = as.factor(d$zariz_3)
d$zariz_4 = as.factor(d$zariz_4)
d$posob_1 = as.factor(d$posob_1)
d$posob_2 = as.factor(d$posob_2)
d$posob_3 = as.factor(d$posob_3)
d$posob_4 = as.factor(d$posob_4)
```

Samotná analýza probíhá pomocí knihovny *mlogit* (Croissant, 2011, 2012). Knihovna se načte a data se připraví do tvaru pro modelování.

```
library(„mlogit“)
md = mlogit.data(d, „Choice“, shape=„wide“, varying=7:22, sep=„_“, id.var=„id“)
```

Příkazu *mlogit.data* se určí, s jakou tabulkou dat se bude pracovat, jaké jméno má proměnná, která obsahuje volbu respondenta, z jakého typu matice se datová sada pro multinomiální logit vytváří. Označení „wide“ znamená, že v původní tabulce je každé pozorování na jednom řádku a atributy všech alternativ jsou ve sloupcích za sebou. V proměnné *varying* je uloženo, že atributy alternativ se nachází na sedmém až dvacátém druhém sloupci tabulky. Názvy těchto atributů se skládají ze samotného názvu atributu a označení příslušné alternativy, které jsou od sebe odděleny znakem, uloženým v proměnné *sep*. V proměnné *id.var* se zadává název sloupce tabulky, který obsahuje číslo respondenta. Příkaz má samozřejmě ještě celou řadu dalších nastavení.

Do proměnné *md* dostaneme data v podobě vhodné pro analýzu. Výsledkem je tabulka dat ve formátu „long“. Tentokrát je na každém řádku jedna alternativa. Sloupec *id* obsahuje číslo respondenta, za kterým následují sociodemografické údaje. Sloupec *chid* obsahuje číslo scénáře, respektive pozorované volby a sloupec *alt* nese číslo alternativy, pro kterou jsou ve sloupcích *posob*, *zariz*, *vzdal* a *najem* hodnoty atributů. Sloupec *Choice* obsahuje informaci o tom, která alternativa byla vybrána.

## 5.4 Základní model

Model se skládá z jedné užitkové funkce, která je stejná pro všechny alternativy. Alternativy nejsou pojmenované, a proto nemají žádnou alternativně specifickou složku<sup>4</sup>. Vzhledem k malému počtu respondentů nejsou do modelu zahrnuty ani sociodemografické údaje,

<sup>4</sup> To, zda pořadí alternativy nemá vliv na pravděpodobnost výběru lze testovat přidáním alternativně specifické konstanty. Tyto konstanty však pro naše data nejsou statisticky významně rozdílné od nuly, a proto nejsou do modelu zahrnuty.

kteří slouží k popisu výběrového souboru. Pro alternativu  $i$  by mohla deterministická část vypadat takto:

$$V_i = \beta_{najem} najem_i + \beta_{posob} posob_i + \beta_{vzda1} vzdal_i + \beta_{zariz} zariz_i \quad (28)$$

V prostředí  $R$  vypadá zadání modelu následovně. Nula v rovnici znamená, že se v modelu nemají použít alternativně specifické konstanty.

```
m1 = mlogit(Choice ~ najem + posob + vzdal + zariz + 0, md)
summary(m1)
```

Výpis modelu obsahuje především tabulku odhadů parametrů modelu, chyby těchto odhadů a výsledky t-testů rozdílů od nuly. Pro kategoriální proměnné (*posob*, *zariz*) se odhadují parametry pro jednotlivé hodnoty proměnných.

```
Call:
mlogit(formula = Choice ~ najem + posob + vzdal + zariz + 0,
       data = md, method = „nr“, print.level = 0)

Frequencies of alternatives:
      1      2      3      4
0.23077 0.25214 0.27778 0.23932

nr method
6 iterations, 0h:0m:0s
g'(-H)^-1g = 0.000376
successive function values within tolerance limits

Coefficients :
      Estimate Std. Error t-value Pr(>|t|)
najem  -0.00082395  0.00008106 -10.1647 < 2.2e-16 ***
posob2 -1.86536549  0.26248559  -7.1065 1.190e-12 ***
posob3 -3.69928484  0.37332388  -9.9090 < 2.2e-16 ***
vzda1  -0.05080133  0.01118997  -4.5399 5.628e-06 ***
zariz1 -0.19582217  0.26253155  -0.7459  0.45573
zariz2  0.47133194  0.24434014   1.9290  0.05373 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log-Likelihood: -153.69
```

Nespočítá se ovšem McFaddenovo  $R^2$ , to se vypíše pouze pro modely s alternativně specifickými konstantami. Výpis ale vrátí *log-likelihood* modelu, který nám pro výpočet  $R^2$  stačí. Musíme ovšem spočítat *log-likelihood* nulového modelu. V našem případě není mezi jednotlivými alternativami, jak jsou respondentům nabízeny, žádný apriorní rozdíl. Alternativy se liší pouze ve svých atributech. Proto můžeme předpokládat nulový model, kde má každá alternativa stejnou pravděpodobnost zvolení. Alternativy jsou čtyři

a pravděpodobnost zvolení některé z nich je  $\frac{1}{4}$ . Logaritmizovanou věrohodnost nulového modelu dostaneme jednoduše jako logaritmus pravděpodobnosti výběru alternativy (jedna čtvrtina), násobený počtem pozorování. *Log-likelihood* nulového modelu zapíšeme do proměnné *LL0* a jeho hodnota je -324,39.

```
LL0 = log(1/4)*nrow(d)
cat(„R2(m1) = „,1-as.numeric(m1$logLik)/LL0,“\n“)
AIC(m1)
```

*Log-likelihood* modelu je -153,69.  $R^2$  pak vychází 0,526. Zároveň můžeme vypočítat například Akaikovo informační kritérium, které pro tento model vychází 165,7.

Model předpokládá, že chyby jednotlivých pozorování budou nezávislé. My ale víme, že každému respondentovi byly předloženy tři scénáře. Individuální rozdíly mezi respondenty mohou vstupovat do rozdílů v percepci jednotlivých atributů (jejich váhy při rozhodování). Tyto rozdíly v modelu vstupují do chyby modelu. Chyby všech pozorování pro jednoho respondenta tedy mohou být na sobě závislé. Vzniká tak riziko špatného odhadu parametrů modelu, což lze odstranit použitím smíšeného multinomiálního modelu pro panelová data (více též v kapitole 2.5).

## 5.5 Smíšený model

Smíšený model pro panelová data znamená, že se budou odhadovat některé parametry s ohledem na každého respondenta. Bude se předpokládat, že tyto odhadované parametry modelu budou variovat mezi respondenty v rámci nějakého typu rozdělení a s určitým rozptylem. Tuto varianci bude model kromě průměru odhadovat také. Které parametry budou v modelu vystupovat jako náhodné proměnné, je třeba vyzkoušet. Vyzkoušením aplikace náhodné proměnné na všechny parametry se ukázalo, že pro model je nejvýhodnější jako náhodné uvažovat parametry pro nájem a počty osob, zatímco parametry pro vzdálenost od školy a zařízení pokoje zůstávají bez náhodné komponenty. Proměnná počtů osob je nominální a v modelu se vyskytuje jako dvě dummy proměnné. Proto se pro ni odhadují dva parametry (*posob2* a *posob3*).

```
m2=mlogit(Choice ~ najem + posob + vzdal + zariz + 0, md, panel=T,
rpar=c(najem="n",posob2="n",posob3="n"), R=1000, start=c(c(m1$coefficients),0,0,0))
summary(m2)
```

Zadání modelu se nyní změnilo. Určíme v něm, že se má jednat o panelová data (*panel=T*). Dále určíme, které proměnné má procedura považovat za náhodné a jaký typ rozdělení má u nich očekávat. Označení „n“ znamená normální rozdělení, ale k dispozici jsou i další typy rozdělení náhodných veličin, které lze zvolit. K odhadu smíšeného modelu se používá metoda MSLE (*Maximum Simulated Likelihood Estimation*), kde se pro stanovení

věrohodnosti modelu použijí hodnoty náhodných veličin simulovaných podle jejich rozdělení (Train, 2001; Hole, 2007). Proto se zadává počet těchto simulací hodnot v jednom kroku odhadu parametrů (v tomto případě  $R=1000$ ). Čím je počet simulací větší, tím lépe jsou parametry náhodných proměnných odhadovány, ale zároveň tím také roste doba potřebná pro odhad modelu. Nakonec je třeba zadat výchozí velikosti odhadovaných parametrů. Bez nich model často není schopen najít řešení. Lze předpokládat, že průměrné velikosti parametrů smíšeného modelu budou podobné, jako u předchozí verze modelu. Proto jsou tyto počáteční hodnoty průměrů parametrů vzaty z odhadu předchozího modelu a směrodatné odchylky tří náhodných parametrů jsou na začátku nastaveny na nulu.

```
Call:
mlogit(formula = Choice ~ najem + posob + vzdal + zariz + 0,
        data = md, start = c(c(m1$coefficients), 0, 0, 0), rpar = c(najem = „n“,
        posob3 = „n“, posob2 = „n“), R = 1000, panel = T)

Frequencies of alternatives:
      1      2      3      4
0.23077 0.25214 0.27778 0.23932

bfgs method
28 iterations, 0h:0m:18s
g'(-H)^-1g = 7.39E-07
gradient close to zero

Coefficients :
      Estimate Std. Error t-value Pr(>|t|)
najem    -0.00194314  0.00049790  -3.9027  9.514e-05 ***
posob2   -3.63790963  0.97069034  -3.7478  0.0001784 ***
posob3   -8.53685617  2.07048992  -4.1231  3.738e-05 ***
vzdal    -0.10525769  0.02999990  -3.5086  0.0004505 ***
zariz1   -0.69942864  0.48634812  -1.4381  0.1503990
zariz2    1.18317496  0.51317117   2.3056  0.0211322 *
sd.najem  0.00090143  0.00027031   3.3348  0.0008537 ***
sd.posob2 3.07778107  1.09653424   2.8068  0.0050032 **
sd.posob3 -4.00494770  1.36125473  -2.9421  0.0032599 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log-Likelihood: -132.27

random coefficients
      Min.      1st Qu.      Median      Mean      3rd Qu. Max.
najem  -Inf    -0.002551148  -0.001943142  -0.001943142  -0.001335136  Inf
posob3  -Inf   -11.238152347  -8.536856171  -8.536856171  -5.835559994  Inf
posob2  -Inf    -5.713841417  -3.637909633  -3.637909633  -1.561977848  Inf
```

Výpis nyní ukazuje navíc odhady směrodatných odchylek pro náhodné parametry a také popis jejich rozdělení. Můžeme si všimnout, že směrodatná odchylka pro náhodnou komponentu počtu osob tří (*sd.posob3*) je záporná. To je zjevný nesmysl, protože směrodatná odchylka nemůže nabývat záporných hodnot. Záporné hodnoty u směrodatných odchylek

náhodných parametrů jsou vedlejším efektem *MSLE* algoritmu odhadu. U symetrických rozdělení, jako je zde, nemá znaménko u odhadu směrodatné odchylky žádný efekt a lze odhad interpretovat jako kladný. Přehledné informace o náhodných proměnných v modelu můžeme získat také procedurou „*rpar(m2)*“ nebo si nechat vykreslit odhadnuté rozdělení náhodné proměnné pomocí příkazu „*plot(rpar(m2,„najem“))*“. Pro srovnání kvality smíšeného modelu s prvním modelem můžeme vypočítat  $R^2$  a informační kritérium.

```
cat(„R2(m2) = „,1-as.numeric(m2$logLik)/LL0,“\n“)
AIC(m2)
```

*Log-likelihood* modelu stoupl na -132,27 a  $R^2$  je nyní 0,592. Akaikovo informační kritérium pro tento model vychází 150,3. Je tedy u tohoto modelu menší ztráta informace oproti modelu verze 1. Můžeme rovněž srovnat první a druhý model pomocí testu poměru věrohodnosti (kap. 4.1).

```
Likelihood ratio test

Model 1: Choice ~ najem + posob + vzdal + zariz + 0
Model 2: Choice ~ najem + posob + vzdal + zariz + 0
#Df LogLik Df Chisq Pr(>Chisq)
1 6 -153.69
2 9 -132.27 3 42.829 2.676e-09 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Test poměru věrohodnosti ukazuje významný rozdíl ve věrohodnosti obou modelů.

Odhadnuté parametry smíšeného modelu jsou v absolutní hodnotě větší, než u první verze modelu. To je proto, že část chyby modelu je nyní „vysvětlena“ náhodnými parametry. Jestli jsou oba modely skutečně podstatně rozdílné, uvidíme při kvantifikaci parametrů.

## 5.6 Kvantifikace parametrů

V modelech používáme cenu měsíčního nájmu za jednu osobu. Vůči parametru této proměnné budeme posuzovat ostatní parametry. Pro výpočet chyby odhadu budeme kromě odhadu parametrů také potřebovat standardní chyby odhadu. Ty jsou uvedeny ve výpisu v téže tabulce. Odhad chyby kvantifikovaných parametrů můžeme dále zpřesnit započtením kovariancí mezi standardními chybami parametrů (příkaz *vcov* v prostředí *R*). Celou záležitost za nás udělá v prostředí *R* funkce popsaná v kapitole 4.2, která vrátí tabulku kvantifikovaných parametrů zadaného modelu a názvu referenčního parametru, což je v našem případě pro oba modely „*najem*“. Je potřeba podotknout, že atribut pro nájem není měřen v korunách, ale v korunách za měsíc (jedná se o měsíční nájem).

```
kvantif(m1, "najem")
kvantif(m2, "najem")
```

První funkce vrátí výpis pro model 1:

```

           val           se           cov           kva1           kse
najem    -0.0008239478  8.106001e-05  6.570725e-09  1.0000  1.862645e-09
posob2   -1.8653654932  2.624856e-01  9.208715e-06  2263.9364  2.994602e+02
posob3   -3.6992848437  3.733239e-01  1.626936e-05  4489.7076  4.303490e+02
vzda1    -0.0508013290  1.118997e-02  2.363503e-07  61.6560  1.335306e+01
zariz1   -0.1958221708  2.625316e-01  1.396062e-06  237.6633  3.179497e+02
zariz2    0.4713319357  2.443401e-01 -2.602699e-06 -572.0410  2.944856e+02

```

Druhá funkce zase výpis pro druhý model:

```

           val           se           cov           kva1           kse
najem    -0.001943142  0.0004979006  2.479051e-07  1.0000000  5.268356e-09
posob2   -3.637909633  0.9706903398  4.186998e-04  1872.1793691  2.538952e+02
posob3   -8.536856171  2.0704899239  9.172867e-04  4393.3268314  5.176945e+02
vzda1    -0.105257690  0.0299999046  1.056639e-05  54.1688209  1.130636e+01
zariz1   -0.699428637  0.4863481184  8.821972e-05  359.9473316  2.330911e+02
zariz2    1.183174956  0.5131711724 -1.265213e-04 -608.8979568  2.308277e+02
sd.najem  0.000901431  0.0002703139 -1.081035e-07 -0.4639039  8.317491e-02
sd.posob2 3.077781069  1.0965342358 -4.482982e-04 -1583.9201086  3.271832e+02
sd.posob3 -4.004947704  1.3612547277  5.116169e-04  2061.0683668  4.594820e+02

```

Výpis obsahuje kromě původních hodnot odhadnutých parametrů také standardní chybu odhadu a kovarianci s referenčním parametrem. To jsou všechny údaje, potřebné k výpočtu kvantifikovaných parametrů a standardní chyby této kvantifikace.

Výsledky pro oba modely s vypočtenými kvantifikacemi parametrů (ochotou platit nebo ochotou akceptovat) jsou shrnuty v následující přehledné tabulce, kde nejsou zahrnuty parametry náhodných proměnných u druhého modelu.

Tabulka 6: Kvantifikace parametrů modelů volby pronájmu

para- metr	Model 1					Model 2				
	hodno- ta	std. chyba	kovari- ance	kvantifikace		hodno- ta	std. chyba	kovari- ance	kvantifikace	
				odhad	chyba				odhad	chyba
<i>najem</i>	-0.0008	8.11E-05	6.57E-09	-	-	-0.00194	0.00050	2.48E-07	-	-
<i>posob2</i>	-1.8653	0.26248	9.21E-06	2264	299	-3.63791	0.97069	4.19E-04	1872	254
<i>posob3</i>	-3.6992	0.37332	1.63E-05	4490	430	-8.53686	2.07049	9.17E-04	4393	518
<i>vzda1</i>	-0.0508	0.01119	2.36E-07	62	13	-0.10526	0.03000	1.06E-05	54	11
<i>zariz1</i>	-0.19582	0.26253	1.40E-06	238	318	-0.69943	0.48635	8.82E-05	360	233
<i>zariz2</i>	0.47133	0.24434	-2.60E-06	-572	294	1.18317	0.51317	-1.27E-04	-609	231

Na první pohled je vidět, že zatímco odhadnuté parametry se u obou modelů dosti liší, odhad kvantifikace parametrů je v rámci chyby srovnatelný.

Přesto budeme interpretovat pouze smíšený model (model 2), protože je věcně správnější a o něco lépe odpovídá datům.

Největší vliv na rozhodování respondentů má z nefinančních atributů počet osob na pokoji. Respondenti by byli ochotni akceptovat jednoho spolubydlícího za snížení nájmu o 1872 Kč měsíčně. Pro dva spolubydlící je to 4393 Kč měsíčně, což je o něco více než dvojnásobek. Nové zařízení pokoje si respondenti cení kladně a jsou za ně ochotni zaplatit 609 Kč měsíčně proti nezařízenému pokoji. Naopak, pokud by pokoj byl zařízený starým nábytkem, budou to respondenti vnímat přibližně stejně jako nezařízený pokoj. Rozdíl je statisticky nevýznamný. Vzdálenost od školy hraje překvapivě významnou roli. Proměnná vzdálenost je dána v minutách a proto vychází, že každé minuty potřebné k cestě do školy si respondenti cení na 54 Kč měsíčně. V experimentu byly použity dvě úrovně tohoto atributu, a to 10 minut a 30 minut. Vysokou hodnotu času, stráveného na cestě do školy lze přičíst skutečnosti, že zatímco při desetiminutové cestě do školy půjde respondent pěšky, pro cestu, která trvá 30 minut, bude patrně nucen využít veřejnou dopravu nebo nějaký individuální dopravní prostředek. Studenti jsou ochotni platit za bydlení vzdálené 10 minut od školy proti bydlení 30 minut od školy navíc 1080 Kč měsíčně. To vypadá jako celkem překvapivá ochota platit za bydlení v blízkosti školy. Rozdíl v cestovním čase do školy je mezi objema úrovněmi atributu 20 minut. Kdyby student absolvoval tuto trasu 2x za školní den (do školy a ze školy) a školní den by nastal 20x za měsíc, znamenal by rozdíl mezi úrovněmi atributu celkem 800 minut měsíčně. Pak by hodnota jedné minuty cestovního času byla 1,35 Kč. To je 81 Kč za hodinu. Hodnota cestovního času na cestách po Brně a okolí (viz kapitola 6) vychází na 2,1 Kč/min respektive 146 Kč/h. V této perspektivě se zdá částka na minutu vzdálenosti od školy naopak celkem nízká. Na rozdíl od průměrných Brňanů studenti tolik nespěchají.

## 5.7 Shrnutí

V této případové studii jsme vyzkoušeli design průzkumu, který se zabýval preferencemi studentů při hledání studentského bydlení. Šlo o průzkum vyjádřených preferencí, který proběhl na vzorku 78 respondentů vybraných nahodilým výběrem.

Pro výsledky jsme formulovali jednoduchý model s nepojmenovanými alternativami, který jsme po té upravili do podoby smíšeného modelu pro panelová data. Výsledné efekty jsme následně kvantifikovali, což umožnilo jejich vzájemné srovnání a dalo nám to představu o absolutní velikosti těchto efektů.

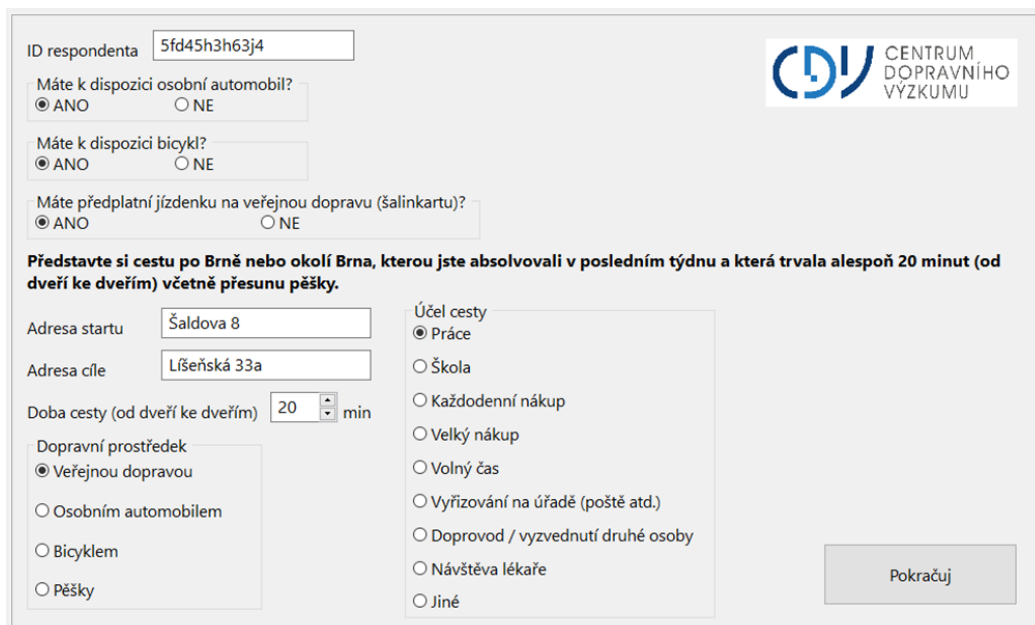
Z výsledků vyplývá, že studenti jsou ochotni akceptovat spolubydlícího za slevu 1872 Kč na měsíčním nájmu. Dále jsou ochotni platit za nově zařízený pokoj o 609 Kč měsíčně navíc proti nezařízenému pokoji, ale i pokoji zařízenému starým nábytkem. Dále jsou ochotni platit za bydlení vzdálené 10 minut od školy proti bydlení 30 minut od školy 1080 Kč měsíčně navíc.

## 6. Případová studie: volba dopravního prostředku na cestu po brněnské metropolitní oblasti (BMO)

Tento SP experiment proběhl na podzim 2016 a jaře 2017 jako součást sociologického výzkumu Dopravní chování obyvatel města Brna a Brněnské metropolitní oblasti (FOCUS, CDV, 2017). Cílem této dílčí části bylo zjistit, jak různé atributy způsobů dopravy ovlivňují jejich výběr.

V rámci dotazování si měli respondenti vzpomenout na nedávnou cestu v rámci BMO a specifikovat cestovní čas, použitý dopravní prostředek a účel cesty. Mělo jít o cestu, která (bez ohledu na zvolený dopravní mód) měla trvat alespoň 20 minut. Tímto způsobem se vyřadily velmi krátké cesty, kde by volba dopravního prostředku nedávala smysl.

V elektronickém formuláři respondenti zadávali údaje o této cestě.



ID respondenta

Máte k dispozici osobní automobil?  
 ANO  NE

Máte k dispozici bicykl?  
 ANO  NE

Máte předplatní jízdenku na veřejnou dopravu (šalinkartu)?  
 ANO  NE

**Představte si cestu po Brně nebo okolí Brna, kterou jste absolvovali v posledním týdnu a která trvala alespoň 20 minut (od dveří ke dveřím) včetně přesunu pěšky.**

Adresa startu

Adresa cíle

Doba cesty (od dveří ke dveřím)  min

Dopravní prostředek  
 Veřejnou dopravou  
 Osobním automobilem  
 Bicyklem  
 Pěšky

Účel cesty  
 Práce  
 Škola  
 Každodenní nákup  
 Velký nákup  
 Volný čas  
 Vyřizování na úřadě (poště atd.)  
 Doprovod / vyzvednutí druhé osoby  
 Návštěva lékaře  
 Jiné

Obrázek 19: Úvod do experimentu

Po té respondenti dostali za úkol vybrat dopravní prostředek pro tuto cestu v rámci simulovaného scénáře.



Nyní si představte, že byste měli absolvovat tuto cestu. V úvahu pro Vás nyní přichází možnosti dopravy s trochu odlišnými vlastnostmi. Který z alternativních způsobů přepravy byste si vybral/a?

**Shrnutí cesty:**

Cesta z Šaldova 8 do Líšeňská 33a.  
Účel cesty byl Práce  
Doba trvání cesty byla 20 min.  
Cestovalo se Veřejnou dopravou

1. z 8.

**Osobní automobil**

Celkový čas cesty 15 min.  
Parkuji zadarmo.

**Pěšky**

Celkový čas cesty 65 min.  
Cesta NENÍ nijak příjemná a bezpečná.

**Bicykl**

Celkový čas cesty 15 min.  
Většinu cesty jedu po normální silnici.

NEMÁM k dispozici zabezpečené parkování kola.

**Veřejná doprava**

Celkový čas cesty 25 min.  
Cestu mám předplacenou.  
Jedu s jedním přestupem.

Obrázek 20: Samotný design experimentu

V tomto scénáři byly čtyři alternativní dopravní módy (automobil, chůze, jízda na kole, veřejná doprava). Úkolem respondenta vybrat tu z nabízených alternativ, kterou by pro danou cestu nejspíš použil.

Atributy alternativ se přepočítaly podle údajů, které respondent sdělil o skutečně absolvované cestě. Využil se přitom především údaj o cestovním času a použitém dopravním módu. Pokud byla tato skutečná cesta provedena pěšky, byly základní časy ostatními módy zmenšeny a naopak, pokud byla cesta vykonána jiným módem, byl základní čas pro pěší cestovní mód navýšen. Dále, v případě, že respondent neměl k dispozici pro tuto cestu automobil nebo jízdní kolo, nebyly mu tyto dopravní prostředky nabízeny. Rovněž v případě, že respondent vlastní časovou jízdenku na veřejnou dopravu, zohlednila se tato skutečnost ve scénáři. Tím se dosáhlo určité realističnosti v předložených scénářích.

## 6.1 Příprava experimentu

Scénář, alternativní dopravní módy a atributy alternativ byly specifikovány ve spolupráci a na základě požadavků zadavatele studie – magistrátu města Brna. Základní cestovní časy, cena za dopravu i další atributy variovaly podle faktoriálního designu.

Experiment byl personalizovaný pro každého respondenta. Byla provedena pilotáž na vzorku 50 osob. Při pilotáži byly nastaveny úrovně cestovního času v SP experimentu.

## 6.2 Výběrový soubor

Skript pro R a tabulka dat k tomuto experimentu jsou soubory „*model\_modechoice.R*“ a „*data\_modechoice.dat*“ (<https://www.cdv.cz/vyzkum-diskretnich-voleb/>). Výběrový soubor tvořilo 2436 respondentů, z toho 1301 jich byli obyvatelé Brna, 1135 obyvatelé brněnského zázemí (další obce BMO kromě Brna). Pro výběr byl použit stratifikovaný vícestupňový náhodný výběr. Výsledný výběrový soubor byl reprezentativní z hlediska pohlaví, věku, městské části (v případě Brna) nebo velikosti obce (v případě ostatních obcí Brněnské metropolitní oblasti - BMO). Sociodemografická data výběrového souboru jsou v tabulce.

Tabulka 7: Sociodemografie respondentů

	n	%
Celkem	2436	100
Pohlaví		
muži	1120	46
ženy	1316	54
Věk		
15 - 24 let	241	10
25 - 34 let	322	13
35 - 44 let	519	21
45 - 54 let	391	16
55 - 64 let	373	15
65 a více let	487	20
neuveдено	103	4
Vzdělání		
ZŠ	158	6
vyučen-a	594	24
SŠ s maturitou	1023	42
VŠ	661	27
Ekonomická aktivita		
samostatně výdělečně činný	273	11
zaměstnanec	1155	47
nezaměstnaný	43	2
důchodce	552	23
neuveдено	413	17

Při analýze volby dopravního módu nebudeme s údaji o sociodemografii respondentů pracovat a proto není obsažena v datovém souboru k tomuto příkladu. Cílem příkladu je ukázat možnosti kvantifikace proměnných, a to především stanovení hodnoty času.

## 6.3 Příprava dat

Pro analýzu *SP* experimentu na volbu dopravního módu pomocí knihovny *mlogit* v prostředí *R* bude potřeba data nejprve poněkud upravit. Budeme se muset vypořádat s tím, že v rámci *SP* experimentu nebyly administrovány alternativy osobního automobilu a jízdního kola, pokud respondent těmito dopravními prostředky nedisponoval. Druhým problémem je fakt, že pouze cestovní čas byl atribut, který obsahovaly všechny typy alternativ. Ostatní atributy se vyskytovaly pouze u příslušných alternativ.

```
library(„mlogit“)
rm(list=ls()) # smaže proměnné

d=read.table(„data_brno_R2.dat“, header = T)
```

V prostředí *R* nejprve načteme datovou matici. Tato matice obsahuje proměnné, které označují, zda měl respondent k dispozici osobní automobil (*cavl*) a jízdní kolo (*bavl*) a zda vlastní předplatní jízdenku na městskou hromadnou dopravu v Brně (*ptcard*). Následuje proměnná, která obsahuje kategorizaci účelu cesty (*purpose*). Účely mohou nabývat těchto hodnot:

- 0 – cesta do zaměstnání
- 1 – cesta do školy
- 2 – každodenní nákup
- 3 – velký nákup
- 4 – cesta za volnočasovou aktivitou
- 5 – vyřizování na poště nebo úřadech
- 6 – doprovod jiné osoby
- 7 – návštěva lékaře
- 8 – ostatní

Další proměnnou je kategorizace dopravního prostředku, který respondent ve skutečnosti použil pro tuto cestu (*lastmode*). Jde o kategorie:

- p – veřejná doprava
- c – osobní automobil
- b – jízdní kolo
- w – cestování pěšky

Stejnou kategorizaci mají také jednotlivé alternativy a proměnná alternativy zvolené respondentem v rámci *SP* experimentu (*Choice*).

Jako atributy alternativních způsobů dopravy byly v rámci experimentu respondentům prezentovány tyto:

- Všechny alternativy
  - cestovní čas pro jednotlivé dopravní módy ( $tt_p, tt_c, tt_b, tt_w$ ) je v minutách;

- Veřejná doprava
  - cena za veřejnou dopravu (*cost*), která nabývala hodnot 20 a 25 Kč za cestu. V případě, že respondent deklaroval vlastnictví časové jízdenky, nastaví se cena automaticky na nulu – za tuto konkrétní cestu respondent neplatil;
  - počet přestupů ve veřejné dopravě (*trans*), byl nastaven na žádný nebo jeden přestup;
- Automobilová doprava
  - cena za parkování osobního automobilu (*pcost*) nabývá dvou hodnot: „parkujete zadarmo“ nebo „za parkování platíte 30 Kč/h“;
- Cyklistická doprava
  - parametr cyklistické infrastruktury (*infrab*) nabývá hodnot „Většinu cesty jedu po normální silnici.“ nebo „Většinu cesty jedu po vyhrazeném cyklopruhu.“;
  - možnost parkování kol (*parkb*) „NEMÁM k dispozici zabezpečené parkování kola.“; „MÁM k dispozici zabezpečené parkování kola.“;
- Pěší doprava
  - kvalita pěší trasy (*infrac*) nabývá hodnot „Cesta NENÍ nijak příjemná a bezpečná.“; „Cesta JE příjemná a bezpečná.“ a násobí se cestovním časem pěšky.

Poslední proměnnou v tabulce je indikátor toho, zda řádek náleží obyvateli Brna nebo brněnského zázemí (*Brno*).

Nyní je potřeba převést data do tvaru pro multinomiální logistický model. Pouze proměnná cestovního času (*tt*) nabývá alternativně specifických hodnot (nachází se ve sloupcích 6 až 9). Ostatní proměnné náleží vždy jedné alternativě.

```
md=mllogit.data(d, "Choice", shape="wide", varying=6:9, sep="_")
```

V řádcích máme jednotlivé alternativy každého výběru. Alternativ, které nebyly respondentovi administrovány, pokud neměl k dispozici automobil nebo kolo, se zbavíme ponecháním jen příslušných řádků.

```
md=md[(md$cav==1) | (md$alt!="c"), ]
md=md[(md$bav==1) | (md$alt!="b"), ]
```

Abychom vyřešili problém s proměnnými, které se vyskytují jen u svých alternativ, potřebujeme, aby u alternativ, ve kterých se nevyskytují, nabývaly tyto proměnné nulové hodnoty. Proto si nejprve vytvoříme indikátory, které nabývají hodnoty 1, pokud řádek náleží příslušné alternativě. Těmi budeme násobit příslušné proměnné, specifické pro alternativy.

```
md$b=ifelse(md$alt=="b",1,0)
md$c=ifelse(md$alt=="c",1,0)
md$p=ifelse(md$alt=="p",1,0)
md$w=ifelse(md$alt=="w",1,0)
```

Důležitou informací, která může značně ovlivnit volbu respondenta v *SP* experimentu, je to, jaký způsob dopravy respondent zvolil pro tuto cestu ve skutečnosti. Vytvoříme si proto alternativně specifickou proměnnou, která bude nabývat hodnoty jedna, pokud řádek alternativy odpovídá užitému způsobu dopravy.

```
md$last = ifelse(md$alt==md$last,1,0)
```

## 6.4 Model pro celková data

Model se skládá z pojmenovaných alternativ a obsahuje alternativně specifické proměnné, které se, kromě cestovního času, vyskytují pouze v příslušných typech alternativ. Cestovní čas se jediný vyskytuje ve všech alternativách a v tomto modelu bude mít pro všechny alternativy stejný parametr  $\beta_{tt}$ .

$$V_p = ASC_p + \beta_{tt} tt_p + \beta_{lastp} last_p + \beta_{cost} cost + \beta_{trans} trans$$

$$V_c = ASC_c + \beta_{tt} tt_c + \beta_{lastc} lastc + \beta_{pcost} pcost$$

$$V_b = ASC_b + \beta_{tt} tt_b + \beta_{lastb} lastb + \beta_{infrab} infrab + V_b + \beta_{parkb} parkb$$

$$V_w = ASC_w + \beta_{tt} tt_w + \beta_{lastw} lastw + \beta_{infracw} infracw$$

V prostředí *R* lze model zapsat následovně.

```
mn = mlogit(Choice ~ I(cost*p) + I(pcost*c) + I(trans*p) + I(parkb*b) + I(infracw*w)
+ I(infracb*b) + tt | 1 | last , md)
```

Alternativně specifické proměnné jsou násobeny indikátorem příslušné alternativy. Proto nabývají nějaké hodnoty pouze u příslušné alternativy. Jejich popis se nachází v první sekci, kde jsou alternativně specifické proměnné asociovány s nesespecifickými parametry. Zde se nachází také proměnná cestovního času. Ve druhé sekci zůstávají pouze alternativně specifické konstanty. Proměnná skutečně použitého dopravního módu (*last*) je v sekci alternativně specifických proměnných s alternativně specifickými parametry, protože preference na základě dříve použitého způsobu dopravy se může lišit v závislosti na tom kterém dopravním módu.

```
Call:
mlogit(formula = Choice ~ I(cost * p) + I(pcost * c) + I(trans *
p) + I(parkb * b) + I(infrac * w) + I(infrac * b) + tt |
1 | last, data = md, method = „nr“, print.level = 0)
```

Frequencies of alternatives:

	b	c	p	w
	0.081281	0.401067	0.495484	0.022167

nr method

8 iterations, 0h:0m:0s

$g'(-H)^{-1}g = 0.00053$

successive function values within tolerance limits

Coefficients :

	Estimate	Std. Error	z-value	Pr(> z )
c:(intercept)	2.3878356	0.1525169	15.6562	< 2.2e-16 ***
p:(intercept)	2.2560588	0.2074307	10.8762	< 2.2e-16 ***
w:(intercept)	1.7861384	0.4248692	4.2040	2.623e-05 ***
I(cost * p)	-0.0538492	0.0056696	-9.4979	< 2.2e-16 ***
I(pcost * c)	-0.0533523	0.0044135	-12.0885	< 2.2e-16 ***
I(trans * p)	-0.3755286	0.1355091	-2.7712	0.005584 **
I(parkb * b)	0.4827711	0.1837658	2.6271	0.008612 **
I(infrac * w)	0.7970064	0.3472239	2.2954	0.021712 *
I(infrac * b)	0.0557018	0.1871357	0.2977	0.765967
tt	-0.1301172	0.0086863	-14.9796	< 2.2e-16 ***
b:last	2.9132401	0.4441764	6.5587	5.426e-11 ***
c:last	1.2335273	0.1781045	6.9259	4.333e-12 ***
p:last	1.2811418	0.1937415	6.6126	3.775e-11 ***
w:last	1.0781543	0.3576207	3.0148	0.002571 **

---

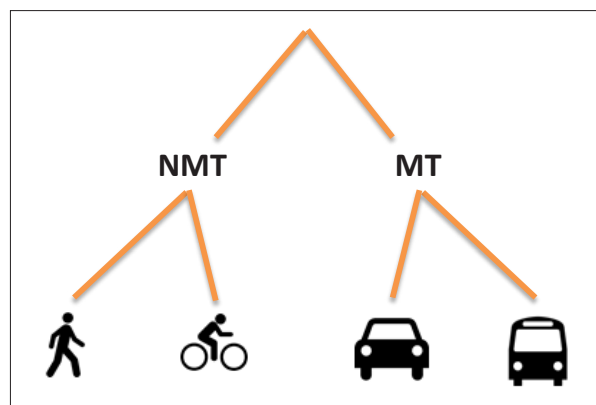
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Log-Likelihood: -1306.2

McFadden  $R^2$ : 0.46528

Likelihood ratio test :  $\text{chisq} = 2273.2$  (p.value = < 2.22e-16)

Model vypočítal přímo McFaddenovo  $R^2$  jako ukazatel kvality modelu. Jeho hodnota 0,46528 pro obecný model je dobrá.



Obrázek 21: Vnořená struktura volby dopravního módu

U volby dopravního módu se často uvažuje o tom, že volba probíhá ve více krocích. Můžeme například uvažovat, že se obyvatelé brněnské metropolitní oblasti nejprve rozhodnou, jestli zvolí motorizovaný (MT) nebo nemotorizovaný (NMT) způsob transportu a po té teprve zvolí konkrétní způsob dopravy. Tato vnořená (*nested*) struktura je na obrázku 21.

Model s vnořenou strukturou (*nested logit*), oddělující motorizované a nemotorizované dopravní módy, bude formulovaný v *R* následovně.

```
n1=mlogit(Choice ~ I(cost*p)+I(pcost*c) + I(trans*p) + I(parkb*b) + I(infraw*w) +
I(infrab*b) + tt | 1 | last, md, nests=list(MT=c(,"c","p"),NMT=c(,"b","w")),un.nest.el=T)
```

Vnořená struktura se do formulace modelu přidá specifikací proměnné *nests*, ve které jsou definovány množiny alternativ při rozhodování první úrovně (MT a NMT). Nastavení proměnné *un.nest.el* na hodnotu TRUE způsobí, že se bude očekávat stejná úroveň korelace pro alternativy ve všech (obou) vnořených množinách (Croissant, 2012; Train & Croissant, 2012). Všechno ostatní může zůstat stejné.

```
Call:
mlogit(formula = Choice ~ I(cost * p) + I(pcost * c) + I(trans *
p) + I(parkb * b) + I(infraw * w) + I(infrab * b) + tt |
1 | last, data = md, nests = list(MT = c(,"c", "p"), NMT = c(,"b",
,"w")), un.nest.el = T)
```

Frequencies of alternatives:

	b	c	p	w
	0.081281	0.401067	0.495484	0.022167

bfgs method

10 iterations, 0h:0m:1s

$g'(-H)^{-1}g = 1.08E-07$

gradient close to zero

Coefficients :

	Estimate	Std. Error	z-value	Pr(> z )	
c:(intercept)	2.2560638	0.1345440	16.7682	< 2.2e-16	***
p:(intercept)	2.1308479	0.1772586	12.0211	< 2.2e-16	***
w:(intercept)	1.1684010	0.4611646	2.5336	0.011290	*
I(cost * p)	-0.0412147	0.0059122	-6.9711	3.144e-12	***
I(pcost * c)	-0.0408355	0.0054055	-7.5544	4.219e-14	***
I(trans * p)	-0.2531649	0.1123218	-2.2539	0.024201	*
I(parkb * b)	0.5375088	0.1723024	3.1196	0.001811	**
I(infraw * w)	0.8466787	0.3736378	2.2660	0.023449	*
I(infrab * b)	0.1037025	0.1721055	0.6026	0.546807	
tt	-0.1063225	0.0121453	-8.7542	< 2.2e-16	***
b:last	2.5812231	0.4570863	5.6471	1.632e-08	***
c:last	1.0026893	0.1808323	5.5449	2.942e-08	***
p:last	1.0609629	0.1949782	5.4414	5.285e-08	***
w:last	1.1543121	0.3908129	2.9536	0.003141	**
iv	0.7187870	0.0850961	8.4468	< 2.2e-16	***
---					

```

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log-Likelihood: -1301.4
McFadden R^2: 0.46725
Likelihood ratio test : chisq = 2282.8 (p.value = < 2.22e-16)

```

Na rozdíl od předchozího modelu se odhadl parametr  $iv$ , který se týká vnořené struktury. Výsledky ukazují, že má tento model nepatrně lepší  $R^2=0,46725$ . Parametr vnoření vychází podle výpisu statisticky významně rozdílný od nuly. To nás ale nezajímá, protože potřebujeme testovat jeho rozdíl od jedné. Pokud by tento parametr byl neodlišitelný od jedné, vnořená struktura by neměla statisticky významný efekt. K otestování můžeme použít t-statistiku.

```
(coef(n1)[,iv'] - 1) / sqrt(vcov(n1)[,iv', ,iv'])
```

Výsledkem testu je hodnota -3,304651. Kritická hodnota t-testu pro 95% interval spolehlivosti je 1,96, takže můžeme zamítnout hypotézu, že je tento parametr roven jedné (Train & Croissant, 2012). Abychom získali jednoznačné vyjádření o kvalitě nového modelu, můžeme využít *Likelihood ratio test*, popsany v kapitole 4.1.

```
lrtest(mn, n1)
```

Pomocí testu poměru věrohodností porovnáme multinomiální model a model s vnořenou strukturou.

```

Likelihood ratio test

Model 1: Choice ~ I(cost * p) + I(pcost * c) + I(trans * p) + I(parkb *
  b) + I(infraw * w) + I(infrab * b) + tt | 1 | last
Model 2: Choice ~ I(cost * p) + I(pcost * c) + I(trans * p) + I(parkb *
  b) + I(infraw * w) + I(infrab * b) + tt | 1 | last
  #Df  LogLik Df  Chisq Pr(>Chisq)
  1  14 -1306.2
  2  15 -1301.4  1  9.6503  0.001893 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Model s vnořenou strukturou má vyšší stupeň volnosti než obyčejný multinomiální model, a to o škálovací parametr vnořené struktury ( $vi$ ). I přesto test ukazuje, že rozdíl mezi oběma modely je významný na jednoprocenní hladině. Podobně dopadnou i srovnání pomocí informačních kritérií prvního, tedy multinomiálního modelu  $mn$  a druhého *nested logit* modelu  $n1$ .



```
cat(„AIC(mn)= „,2*length(mn$coefficients)-as.numeric(mn$logLik),“\n”)
cat(„AIC(nl)= „,2*length(nl$coefficients)-as.numeric(nl$logLik),“\n“)
```

Akaikovo informační kritérium pro multinomiální model je  $AIC(mn)=1334,238$  a pro model s vnořenou strukturou je o něco menší  $AIC(nl)=1333,309$ . Z těchto všech důvodů přijmeme jako lepší model s vnořenou strukturou.

## 6.5 Kvantifikace parametrů modelu pro celková data

Kvantifikaci parametrů a odhad standardní chyby těchto složených proměnných vypočítáme podle vzorců 15 a 16. V prostředí *R* můžeme použít funkci, která vrátí tabulku kvantifikovaných parametrů pro zadaný model a název referenční proměnné. Tato funkce je popsána v kapitole 4.2. Pro náš model je název referenčního parametru ceny „ $l(cost * p)$ “.

Tabulka 8: Kvantifikace parametrů modelu pro celková data

parametr	hodnota	std. chyba	kovariance	kvantifikace	
				hodnota	std. chyba
c:(intercept)	2.2561	0.1345	-0.0001288	-54.7	8.0
p:(intercept)	2.1308	0.1773	-0.0003663	-51.7	7.2
w:(intercept)	1.1684	0.4612	-0.0009572	-28.3	10.5
l(cost * p)	-0.0412	0.0059	0.0000350	-	-
l(pcost * c)	-0.0408	0.0054	0.0000174	1.0	0.1
l(trans * p)	-0.2532	0.1123	0.0001323	6.1	2.7
l(parkb * b)	0.5375	0.1723	0.0001005	-13.0	4.7
l(infraw * w)	0.8467	0.3736	0.0000549	-20.5	9.6
l(infrab * b)	0.1037	0.1721	0.0001534	-2.5	4.2
tt	-0.1063	0.0121	0.0000410	2.6	0.3
b:last	2.5812	0.4571	-0.0003877	-62.6	13.2
c:last	1.0027	0.1808	-0.0003062	-24.3	4.8
p:last	1.0610	0.1950	-0.0002260	-25.7	5.4
w:last	1.1543	0.3908	0.0002864	-28.0	10.7

Kvantifikované parametry se počítají z příslušného parametru a parametru ceny. Parametr ceny je záporný, protože s rostoucí cenou za dopravu klesá užitek takové dopravy. Proto, pokud kvantifikovaný parametr má zápornou velikost, znamená to (poněkud neintuitivně) kladný užitek příslušné proměnné a naopak, kladný kvantifikovaný parametr znamená pro tuto proměnnou její negativní užitek.

Hodnota cestovního času (*tt*) vychází na 2,6 korun za minutu, což je 154,7 Kč/h. Největší efekt na volbu dopravního módu představuje v tomto průzkumu zvyk. Respondent výrazně preferuje dopravní mód, který předtím použil ve skutečnosti. Nejvýraznější je to v případě

jízdního kola ( $b:last$ ), kde je ochota platit za cestu na kole (pokud bylo použito ve skutečnosti) kolem 62,6 Kč. U ostatních dopravních módů je tato preference „ze zvyku“ přibližně poloviční.

Ochota platit za parkování ( $pcost$ ) je přibližně jedna koruna za jednu korunu ceny parkování za hodinu. Z toho vyplývá, že respondenti v průměru uvažovali délku svého placeného parkování na jednu hodinu.

Ochota platit za snížení počtu přestupů ( $(trans * p)$ ) vychází na 6,1 Kč na přestup. Ochota platit za možnost bezpečného parkování kola v cíli cesty ( $(parkb * b)$ ) si respondenti cení na 13 Kč. Cestování po vyhrazeném cyklopruhu ( $(infrab * b)$ ) nevychází v celkovém modelu významná. Možnost cestovat pěšky bezpečnou a příjemnou cestou ( $(infrac * w)$ ) si respondenti cení na 20,5 Kč.

Tento model vyjadřuje preference obyvatel Brna i jeho zázemí. U obou skupin obyvatel se však mohou preference při volbě dopravního módu lišit.

## 6.6 Modely pro Brňany a obyvatele zázemí Brna

Abychom mohli porovnat preference při volbě způsobu dopravy u Brňanů a obyvatel brněnského zázemí, rozdělíme data podle proměnné Brno a na těchto dílčích datech odhadneme parametry předchozího modelu.

```

mdb=md [md$Brno==1, ]
mdz=md [md$Brno!=1, ]

n1B=mlogit(Choice ~ I(cost*p)+I(pcost*c) + I(trans*p) + I(parkb*b) + I(infrac*w) +
I(infrab*b) + tt | 1 | last , mdb,
nests=list(MT=c(„c“, „p“), NMT=c(„b“, „w“)), un.nest.e1=T)
summary(n1B)

n1Z=mlogit(Choice ~ I(cost*p)+I(pcost*c) + I(trans*p) + I(parkb*b) + I(infrac*w) +
I(infrab*b) + tt | 1 | last , mdz,
nests=list(MT=c(„c“, „p“), NMT=c(„b“, „w“)), un.nest.e1=T)
summary(n1Z)

```

Výsledky modelu pro obyvatele Brna pak vypadají následovně:

```

Call:
mlogit(formula = Choice ~ I(cost * p) + I(pcost * c) + I(trans *
p) + I(parkb * b) + I(infrac * w) + I(infrab * b) + tt |
1 | last, data = mdb, nests = list(MT = c(„c“, „p“), NMT = c(„b“,
„w“)), un.nest.e1 = T)

Frequencies of alternatives:
      b      c      p      w
0.076095 0.332052 0.572636 0.019216

bfgs method
8 iterations, 0h:0m:1s
g'(-H)^-1g = 9.27E-07
gradient close to zero

```

```

Coefficients :
      Estimate Std. Error z-value Pr(>|z|)
c:(intercept)  2.7331082  0.2369777 11.5332 < 2.2e-16 ***
p:(intercept)  2.4867908  0.2719367  9.1447 < 2.2e-16 ***
w:(intercept)  1.3587275  0.7698785  1.7649 0.0775874 .
I(cost * p)    -0.0508577  0.0091269 -5.5723 2.515e-08 ***
I(pcost * c)   -0.0441379  0.0076631 -5.7598 8.423e-09 ***
I(trans * p)   -0.0196377  0.1444364 -0.1360 0.8918523
I(parkb * b)   0.7063479  0.2399818  2.9433 0.0032469 **
I(infraw * w)  1.4100417  0.5917445  2.3829 0.0171789 *
I(infrab * b)  0.4611277  0.2449155  1.8828 0.0597270 .
tt             -0.1241566  0.0182568 -6.8006 1.042e-11 ***
b:last         2.8601946  0.9590346  2.9824 0.0028603 **
c:last         0.9680208  0.2635784  3.6726 0.0002401 ***
p:last         1.4032372  0.2979817  4.7091 2.488e-06 ***
w:last         1.2003678  0.6622399  1.8126 0.0698955 .
iv             0.7843803  0.1218239  6.4386 1.205e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log-Likelihood: -653.03
McFadden R^2: 0.47565
Likelihood ratio test : chisq = 1184.7 (p.value = < 2.22e-16)

```

Výsledky pro obyvatele brněnského zázemí jsou zde:

```

Call:
mlogit(formula = Choice ~ I(cost * p) + I(pcost * c) + I(trans *
  p) + I(parkb * b) + I(infraw * w) + I(infrab * b) + tt |
  1 | last, data = mdz, nests = list(MT = c(„c“, „p“), NMT = c(„b“,
  „w“)), un.nest.el = T)

Frequencies of alternatives:
      b      c      p      w
0.076095 0.332052 0.572636 0.019216

bfgs method
10 iterations, 0h:0m:0s
g'(-H)^-1g = 1.31E-06
successive function values within tolerance limits

Coefficients :
      Estimate Std. Error z-value Pr(>|z|)
c:(intercept)  2.7331082  0.2369777 11.5332 < 2.2e-16 ***
p:(intercept)  2.4867908  0.2719367  9.1447 < 2.2e-16 ***
w:(intercept)  1.3587275  0.7698785  1.7649 0.0775874 .
I(cost * p)    -0.0508577  0.0091269 -5.5723 2.515e-08 ***
I(pcost * c)   -0.0441379  0.0076631 -5.7598 8.423e-09 ***
I(trans * p)   -0.0196377  0.1444364 -0.1360 0.8918523
I(parkb * b)   0.7063479  0.2399818  2.9433 0.0032469 **
I(infraw * w)  1.4100417  0.5917445  2.3829 0.0171789 *
I(infrab * b)  0.4611277  0.2449155  1.8828 0.0597270 .
tt             -0.1241566  0.0182568 -6.8006 1.042e-11 ***
b:last         2.8601946  0.9590346  2.9824 0.0028603 **
c:last         0.9680208  0.2635784  3.6726 0.0002401 ***
p:last         1.4032372  0.2979817  4.7091 2.488e-06 ***
w:last         1.2003678  0.6622399  1.8126 0.0698955 .

```

```

iv          0.7843803  0.1218239  6.4386  1.205e-10 ***
---
Signif. codes:  0 '***'  0.001 '**'  0.01 '*'  0.05 '.'  0.1 ' '  1

Log-Likelihood: -632.51
McFadden R^2:  0.45609
Likelihood ratio test : chisq = 1060.8 (p.value = < 2.22e-16)

```

Alternativně specifické konstanty se pro obě skupiny výrazně neliší. Váha přestupu ve veřejné dopravě je u obyvatel Brna statisticky nevýznamná, ale u obyvatel brněnského zázemí je na hranici významnosti. Naopak, možnost bezpečně zaparkovat jízdní kolo je významný faktor pro Brňany ale pro obyvatele zázemí nikoli. Podobně je tomu u významu cyklistické dopravní infrastruktury a příjemné a bezpečné pěší cesty.

## 6.7 Kvantifikace parametrů modelů pro Brňany a obyvatele zázemí Brna

Pro oba modely, odhadnuté v předchozí kapitole byly kvantifikovány parametry modelu. Parametr škály  $vi$  vnořené struktury není uveden, protože pro něj kvantifikace nedává smysl.

Nejdůležitějším údajem je kvantifikace parametru cestovního času, což je hodnota času (viz kapitola 4.2). Ta vychází pro Brňany na  $2,4 \pm 0,4$  Kč/min (146,5 Kč/h) a pro obyvatele zázemí Brna na  $2,8 \pm 0,6$  Kč/min (169,3 Kč/h). Rozdíl mezi hodnotou času obyvatel Brna a obyvatel jeho zázemí je v rámci statistické chyby.

Tabulka 9: Kvantifikace parametrů modelů pro obyvatel Brna a brněnského zázemí

	Model pro Brňany					Model pro obyvatele zázemí Brna				
	hodnota	std. chyba	kovariance	kvantifikace		hodnota	std. chyba	kovariance	kvantifikace	
				hodnota	std. chyba				hodnota	std. chyba
$c:(intercept)$	2.733	0.237	-0.0004	-53.7	9.9	1.994	0.174	-0.0002	-55.1	12.7
$p:(intercept)$	2.487	0.272	-0.0010	-48.9	8.2	2.305	0.444	-0.0009	-63.7	16.6
$w:(intercept)$	1.359	0.770	-0.0027	-26.7	14.0	1.297	0.629	-0.0014	-35.8	17.1
$l(cost * p)$	-0.051	0.009	0.0001	-	-	-0.036	0.008	0.0001	-	-
$l(pcost * c)$	-0.044	0.008	0.0000	0.9	0.2	-0.044	0.009	0.0000	1.2	0.3
$l(trans * p)$	-0.020	0.144	0.0001	0.4	2.8	-0.786	0.410	0.0005	21.7	11.6
$l(parkb * b)$	0.706	0.240	0.0000	-13.9	5.4	0.446	0.296	0.0004	-12.3	9.1
$l(infraw * w)$	1.410	0.592	0.0000	-27.7	12.7	0.310	0.550	0.0002	-8.6	15.4
$l(infrab * b)$	0.461	0.245	0.0002	-9.1	5.2	-0.279	0.304	0.0003	7.7	8.4
$tt$	-0.124	0.018	0.0001	2.4	0.4	-0.102	0.017	0.0001	2.8	0.6
$b:last$	2.860	0.959	-0.0007	-56.2	20.6	2.341	0.541	-0.0007	-64.7	19.3
$c:last$	0.968	0.264	-0.0007	-19.0	5.3	1.160	0.269	-0.0006	-32.0	9.1
$p:last$	1.403	0.298	-0.0006	-27.6	6.8	0.862	0.290	-0.0003	-23.8	9.2
$w:last$	1.200	0.662	0.0010	-23.6	14.4	1.124	0.544	0.0003	-31.1	17.1

Zatímco pro Brňany vychází hodnota přestupu ve veřejné dopravě statisticky nevýznamná, u obyvatel brněnského zázemí je hodnota přestupu  $21,7 \pm 11,6$  Kč. To znamená, že by si obyvatelé zázemí Brna byli ochotni zhruba tuto částku připlatit za snížení počtu přestupů ve veřejné dopravě o jeden přestup. To lze interpretovat tak, že nutnost přestupu ve veřejné dopravě v Brně nepředstavuje velké nepohodlí (průzkum toto nepohodlí nedokázal změřit), ale u obyvatel zázemí představuje přestup ve veřejné dopravě problém, spojený s čekáním nebo rizikem nenavazování spojů.

Možnost bezpečného zaparkování kola si obyvatelé Brna cení na  $13,9 \pm 5,2$  Kč. U obyvatel brněnského zázemí je parametr parkování kola statisticky nevýznamný, protože je zatížen značnou chybou odhadu.

Možnost cestovat pěšky bezpečnou a příjemnou cestou si obyvatelé Brna cení na  $27,7 \pm 12,7$  Kč. Pro obyvatele Brněnského zázemí je tento ukazatel nevýznamný.

## 6.8 Elasticita volby dopravního módu - inkrementální model

V této kapitole se budeme zabývat elasticitou volby dopravního módu. Využijeme k tomu inkrementální model, popsany v kapitole 4.3. Konkrétně nás bude zajímat především to, jak by ovlivnilo zlevnění veřejné dopravy její využívání obyvateli Brna a brněnského okolí.

Abychom se mohli zabývat elasticitou volby dopravního prostředku, potřebujeme znát současný podíl dopravních módů. K tomuto účelu použijeme tyto podíly, zjištěné pomocí cestovních deníků v rámci stejného průzkumu (FOCUS, CDV, 2017), přičemž dopravní módy řidič a spolujezdec byly sloučeny do jedné kategorie.

Tabulka 10: Podíly využívání dopravních módů v BMO

	Obyvatelé Brna		Obyvatelé zázemí Brna	
	všední den	víkend	všední den	víkend
<b>Automobil</b>	38%	51%	61%	66%
<b>VHD</b>	43%	29%	21%	12%
<b>Kolo</b>	1%	0%	1%	1%
<b>Pěšky</b>	18%	20%	17%	21%

Z tabulky vyplývá, že jsou odděleny podíly dopravních módů nejen pro respondenty z Brna a brněnského zázemí, ale také na cesty uskutečněné ve všední den a o víkendu. To je proto, že způsob cestování ve všední den a o víkendu se obvykle dosti liší.

Cena základní jízdenky na 15 minut činí 20 Kč a cena jízdenky na 60 minut 25 Kč<sup>5</sup>. Záleží také na tom, zdali se cestuje ve městě Brně nebo jeho okolí. Budeme chtít znát odpověď na otázku, jak se změní využívání jednotlivých způsobů dopravy, kdyby se snížila cena všech jízdenek o 10 Kč.

Pro inkrementální model je třeba zadat parametr pro cenu za dopravu. Tuto proměnnou hodláme v inkrementálním modelu změnit. Chceme-li se zabývat elasticitou, musíme použít model, který neobsahuje nějaké dodatečné škály, jako jsou koeficienty pro vnořenou strukturu. *Nested logit* modely proto nejsou vhodné, museli bychom u nich parametry přeškálovávat. Procedura *mlogit* nám umožňuje odhadnout neškálované *nested logit* modely, čehož se docílí přidáním parametru „*unscaled=T*“ do procedury. Tak odhadneme stejné modely pro obyvatele Brna a brněnského zázemí, jen s touto drobnou změnou.

```
n1B2=mlogit(Choice ~ I(cost*p)+I(pcost*c) + I(trans*p) + I(parkb*b) +
I(infraw*w) + I(infrab*b) + tt | 1 | last , mdb, nests=list(MT=c(„c“,„p“),
NMT=c(„b“,„w“)), un.nest.el=T, unscaled=T)

n1Z2=mlogit(Choice ~ I(cost*p)+I(pcost*c) + I(trans*p) + I(parkb*b) +
I(infraw*w) + I(infrab*b) + tt| 1 | last , mdz, nests=list(MT=c(„c“,„p“),
NMT=c(„b“,„w“)), un.nest.el=T, unscaled=T)
```

Pro respondenty z Brna dostaneme:

```
Call:
mlogit(formula = Choice ~ I(cost * p) + I(pcost * c) + I(trans *
p) + I(parkb * b) + I(infraw * w) + I(infrab * b) + tt |
1 | last, data = mdb, nests = list(MT = c(„c“, „p“),
NMT = c(„b“, „w“)), un.nest.el = T, unscaled = T)

Frequencies of alternatives:
      b      c      p      w
0.076095 0.332052 0.572636 0.019216

bfgs method
10 iterations, 0h:0m:0s
g'(-H)^-1g = 2.25E-07
gradient close to zero

Coefficients :
      Estimate Std. Error z-value Pr(>|z|)
c:(intercept)  3.4843952   0.5465685   6.3750 1.829e-10 ***
p:(intercept)  3.1703657   0.5525931   5.7373 9.622e-09 ***
w:(intercept)  1.7322396   0.8724965   1.9854 0.047102 *
I(cost * p)    -0.0648375   0.0090011  -7.2033 5.878e-13 ***
I(pcost * c)   -0.0562722   0.0066185  -8.5023 < 2.2e-16 ***
I(trans * p)   -0.0250397   0.1837271  -0.1363 0.891594
I(parkb * b)    0.9005196   0.3365315   2.6759 0.007453 **
I(infraw * w)  1.7976142   0.7978617   2.2530 0.024257 *
I(infrab * b)  0.5878772   0.3345952   1.7570 0.078921 .
tt             -0.1582845   0.0171825  -9.2120 < 2.2e-16 ***
b:last         3.6462269   1.2621151   2.8890 0.003865 **
c:last         1.2341466   0.3086510   3.9985 6.374e-05 ***
p:last         1.7889519   0.3516189   5.0878 3.623e-07 ***
w:last         1.5303386   0.9298606   1.6458 0.099811 .
iv             0.7843877   0.1218245   6.4387 1.205e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log-Likelihood: -653.03
McFadden R^2: 0.47565
Likelihood ratio test : chisq = 1184.7 (p.value = < 2.22e-16)
```

Pro obyvatele zázemí Brna dostaneme:

```

Call:
mlogit(formula = Choice ~ I(cost * p) + I(pcost * c) + I(trans *
p) + I(parkb * b) + I(infraw * w) + I(infrab * b) + tt |
1 | last, data = mdz, nests = list(MT = c(„c“, „p“),
NMT = c(„b“, „w“)), un.nest.el = T, unscaled = T)

Frequencies of alternatives:
      b      c      p      w
0.087225 0.480176 0.407048 0.025551

bfgs method
10 iterations, 0h:0m:0s
g'(-H)^-1g = 1.74E-07
gradient close to zero

Coefficients :
      Estimate Std. Error z-value Pr(>|z|)
c:(intercept)  2.6297840  0.4589322  5.7302 1.003e-08 ***
p:(intercept)  3.0401388  0.7421124  4.0966 4.193e-05 ***
w:(intercept)  1.7098202  0.7482938  2.2850  0.022315 *
I(cost * p)    -0.0477251  0.0089102 -5.3562 8.497e-08 ***
I(pcost * c)   -0.0586044  0.0072171 -8.1203 4.441e-16 ***
I(trans * p)   -1.0371363  0.5199544 -1.9947  0.046079 *
I(parkb * b)    0.5882128  0.4303194  1.3669  0.171650
I(infraw * w)  0.4092233  0.7359104  0.5561  0.578158
I(infrab * b)  -0.3681648  0.4001452 -0.9201  0.357532
tt             -0.1346511  0.0179550 -7.4994 6.417e-14 ***
b:last        3.0864569  0.7867698  3.9229 8.747e-05 ***
c:last        1.5291031  0.3418158  4.4735 7.696e-06 ***
p:last        1.1361551  0.3506510  3.2401  0.001195 **
w:last        1.4824590  0.8141446  1.8209  0.068625 .
iv            0.7583435  0.1313222  5.7747 7.710e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log-Likelihood: -632.51
McFadden R^2: 0.45609
Likelihood ratio test : chisq = 1060.8 (p.value = < 2.22e-16)

```

Můžeme si všimnout, že logaritmy věrohodnosti obou modelů pro Brňany a obou modelů pro obyvatele zázemí Brna jsou přesně stejné, ale parametry těchto modelů se poněkud liší. To je způsobeno tím, že zatímco u první dvojice modelů jsou aplikovány škály vnořené struktury, u druhé dvojice jsou parametry bez těchto škál. Z výsledků obou bezškálových modelů nás budou zajímat pouze parametry ceny „ $I(cost * p)$ “. Parametr ceny je pro Brňany -0,0648375 a pro obyvatele brněnského zázemí -0,0477251.

Do procedury inkrementálního modelu zadáme podíly dopravních módů pro příslušnou skupinu obyvatel a den v týdnu, jak je nalezneme v tabulce 10. Dále proceduře sdělíme, že změna parametru se má týkat druhé alternativy (v našem případě druhého dopravního módu, tedy veřejné dopravy). Jako třetí zapíšeme parametr ceny z příslušného modelu a nakonec hodnotu, o kterou se má cena v inkrementálním modelu změnit.

```

incremental(c(0.38,0.43,0.01,0.18),2,-0.0648375,-10) # Brno - prac. den
incremental(c(0.51,0.29,0.0,0.20),2,-0.0648375,-10) # Brno - víkend
incremental(c(0.61,0.21,0.01,0.17),2,-0.0477251,-10) # zázemí - prac. den
incremental(c(0.66,0.12,0.01,0.21),2,-0.0477251,-10) # zázemí - víkend

```

Tabulka ukazuje, výsledky výpočtů inkrementálních modelů pro obyvatele Brna a brněnského zázemí, pokud by se snížila cena za veřejnou dopravu průměrně o 10 Kč.

Tabulka 11: Podíly dopravních módů obyvatel BMO po zlevnění veřejné dopravy v průměru o 10 Kč

	Obyvatelé Brna		Obyvatelé zázemí Brna	
	všední den	víkend	všední den	víkend
<b>Automobil</b>	27%	40%	54%	61%
<b>VHD</b>	59%	43%	30%	18%
<b>Kolo</b>	1%	0%	1%	1%
<b>Pěšky</b>	13%	16%	15%	20%

Snížení ceny za veřejnou dopravu o 10 Kč by podle inkrementálního modelu znamenalo nárůst podílu využívání právě veřejné dopravy. Bylo by to na úkor především automobilové dopravy, která je jinak využívána nejvíce. Více by se změnilo podíly dopravních módů u obyvatel Brna. Jejich citlivost na cenu veřejné dopravy je vyšší a poptávka elastičtější, než je tomu u obyvatel brněnského zázemí. U obyvatel Brna by taková sleva na veřejné dopravě znamenala teoretický úbytek ve využívání automobilové dopravy o 11 procentních bodů, a to jak ve všední den, tak i o víkendu. Podíly (delších) pěších cest by také byly zasaženy přesunem do veřejné dopravy ale jen úbytkem 5 procentních bodů ve všední den a 4 procentními body o víkendu. Citlivost na cenu ve veřejné dopravě je v tomto modelu větší, než by se dalo očekávat podle průzkumů dopravního chování (*RP dat*). Je to proto, že v rámci experimentálního designu nejsou zahrnuty různé rozdíly mezi reálnými cestami a další možné intervenující proměnné (aktuálně panující počasí a podobně). Z tohoto důvodu se v praxi doporučuje kombinovat SP experimentální data s reálnými daty o dopravním chování (*RP daty*).

## 6.9 Shrnutí

V této případové studii jsme vyzkoušeli model volby dopravního módu s vnořenou strukturou. Vypočítali jsme hodnotu času stráveného v dopravě jak pro všechny respondenty, tak i pro respondenty z Brna a brněnského zázemí. Ukázali jsme také možnosti zhodnocení dalších nefinančních aspektů dopravy, jako jsou přestupy ve veřejné dopravě nebo kvalita prostředí při pěším cestování. Citlivost volby dopravního módu na cenu veřejné dopravy jsme vyzkoušeli inkrementálním modelem.



Všechny zde uvedené modely ukazují vyšší absolutní hodnoty parametrů pro cenu i čas, než jak tomu bývá u analýzy průzkumů dopravního chování (*RP* dat). Zatímco hodnoty parametru pro cestovní čas ve zde prezentovaných modelech oscilovaly od -0,102 do -0,158, z průzkumu dopravního chování v Olomouci 2016-2017 vyplývá parametr pro čas mezi -0,05 až -0,07 pro více zastoupené cesty. To jsou menší hodnoty v absolutních číslech a tak je tomu patrně i u ostatních parametrů. Ve zde prezentovaných modelech jsou odhadnuté parametry většinou v absolutní hodnotě větší než by tomu bylo u modelů získaných na základě *RP* dat a je to způsobeno větší citlivostí, kterou respondenti projevují v hypotetickém experimentu *SP* designu a kterou by neprojevovali v reálných rozhodnutích.

Lišit se také může hodnota času v různých dopravních prostředcích. V takovém případě přesuneme proměnnou cestovního času (*tt*) při formulaci modelu z první do třetí sekce. Tím se stane, že se parametr cestovního času bude odhadovat zvlášť pro každý z dopravních módů.

Dále dává smysl počítat modely zvlášť pro různé účely cesty, kdy parametry modelů mohou vycházet dosti odlišně. Z tohoto důvodu se při dopravním modelování používají pro volbu dopravního módu odlišné modely podle účelu modelované cesty.

V modelech volby dopravního módu se často počítá s různými parametry podle účelu cesty. Obvykle členění cest podle účelů nebývá tak podrobné, jako v tomto případě (devět skupin cest podle účelů) a účely cest se agregují (například na cesty do práce, cesty do školy, nákup, volný čas a ostatní cesty). Můžeme pro jednotlivé účely odhadnout modely separátně a interpretovat jednotlivé parametry modelů s ohledem na účely cestování.

## 7. Závěr

V této knize jsme si představili jeden z přístupů ke studiu rozhodování, a to takový, který se etabloval kromě ekonometrie také v mnoha jiných oborech se snahou ekonometricky podchytit kromě peněžních také nepeněžní přínosy a kde někomu stojí za to nastavit služby tak, aby lépe vyhovovaly jejich uživatelům.

Historicky byla analýza preferencí na odhalených a vyjádřených preferencích vnímána jako zaměnitelné datové zdroje o rozhodování, ale porovnání těchto přístupů vedlo ke zjištění, že mezi výsledky existuje určitá variabilita. Co je silnou stránkou pro *SP* data je slabou stránkou *RP* dat a naopak. (Louviere, Hensher & Swait, 2000a). Použití metody vyjádřených preferencí umožňuje zahrnout také dosud hypotetické alternativy a plně kontrolovat vlastnosti alternativ, které respondent nevyužije. Na druhou stranu s sebou analýza *SP* experimentů přináší riziko spojené s tím, že respondenti vyjadřují pouze hypotetické odpovědi. V takovém případě se dá očekávat větší citlivost respondentů na změnu atributů, než jaká by se u nich dala očekávat v reálných rozhodnutích. Z tohoto důvodu je trendem kombinovat *SP* a *RP* data v jednom modelu diskrétních voleb. To ovšem představuje určitou výzvu.

### 7.1 Kombinace *RP* a *SP* dat

Když získáváme údaje o skutečném rozhodnutí, obvykle od původce rozhodnutí můžeme získat pouze informace o alternativě, kterou zvolil. Pro analýzu ovšem potřebujeme znát také vlastnosti nezvolené alternativy. Můžeme se na ně rovněž zeptat respondenta, ale spolehlivost jeho informací o nezvolených alternativách bude klesat úměrně tomu, jak málo je s těmito alternativami obeznámen. Nakonec to, co se dozvíme, bude spíše to, jak respondent nevyužité alternativy vnímá, než jaké doopravdy jsou. To by teoreticky nemuselo vadit, protože na rozhodování se skutečně podílí právě vnímání charakteristik možných alternativ, než jejich skutečná podoba. Problém je však s tím, že takto získáme nejrozdílnější nekonzistence a výstřednosti, které budou působit potíže při odhadu modelu. Získávání údajů o nevyužitých alternativách také prodlužují dotazování a zatěžují respondenty (Henscher et al., 2007). Adamowicz a kolektiv (1997) například studovali vztah mezi skutečnými a vnímanými atributy alternativ v doméně výběru místa pro rekreaci a zjistili, že ačkoli celkově spolu vnímané a skutečné atributy korelují, v některých případech je jejich korelace velmi nízká.

K řešení tohoto problému se nabízí dva jednoduché přístupy. Vychází z toho, že o každé alternativě nějaké informace máme k dispozici, pouze v případě konkrétního rozhodnutí máme tyto informace pouze o využitých alternativě. Pak můžeme jednoduše pro atributy nepoužitých alternativ použít průměry, respektive mediány atributů těchto alternativ. Nebezpečím tohoto přístupu je skutečnost, že doplňováním chybějících dat středními hodnotami snižujeme variabilitu atributů, ale variabilita rozhodnutí zůstává stejná. Druhou možností je vzít všechny informace o využitých alternativách a (například

náhodně) je přidělit k alternativám stejného typu, ale nevyužitým. Tím se zabrání snížení variability atributů. Oba tyto přístupy mají ovšem daleko k dokonalosti (Henscher et al., 2007).

Z těchto důvodů se, pokud je to možné, přistupuje spíše k modelování všech alternativ. Pro studie z oboru dopravy (například preference druhu dopravy) se nabízí možnost využití dopravních modelů, které umožňují při znalosti místa začátku a cíle cesty vypočítat délku trasy a odhadnout například cestovní časy a ceny za dopravu různými dopravními módy. I zde ovšem zůstává argument, že takto syntetizované údaje o alternativách neodpovídají ani alternativám skutečným, natož tomu, jak tyto alternativy vnímají původci rozhodnutí.

K modelování kombinovaných *RP/SP* dat pak lze použít speciální modely (Ben-Akiva et al., 1994; Hensher et al., 1999; Louviere, Hensher & Swait, 2000a). Tyto modely předpokládají, že se *SP* a *RP* data odhadují se stejnými parametry, ovšem celková škála obou částí modelu se liší. Takové modely jsou vlastně velmi podobné modelům s vnořenou strukturou, kde jsou v jednom „hnízdě“ alternativy z *RP* a ve druhém alternativy z *SP* části dat. Rozdílná škála obou částí modelu je pak škálovací parametr vnořené struktury (Henscher et al., 2007).

## 7.2 Modely diskrétních voleb v dopravních modelech

Dopravní modely jsou složeny ze dvou odlišných částí. Nejprve přijde ke slovu model dopravní poptávky, který (v případě osobní dopravy) modeluje odkud kam a jak se lidé hodlají dopravovat. Jako druhý nastupuje model dopravní nabídky, který pro cestovní poptávku nabídne možnosti cestování – různé trasy dopravní sítě nebo linky veřejné dopravy. Spojení těchto dvou částí dopravního modelu zajišťuje proces přidělování dopravní poptávky na dopravní síť. Výsledkem jsou pak dopravní intenzity na různých úsecích dopravní sítě nebo obsazenost linek veřejné dopravy.

Model dopravní nabídky zahrnuje informace o dopravní síti motorizované, cyklistické i pěší (například délka nebo kapacita silnic, jejich rychlostní limity a povolení vjezdu různými dopravními módy). Dále obsahuje informace o síti veřejné dopravy v modelovaném území, zastávky a linky veřejné dopravy s jejich jízdními řády, cenami za přepravu nebo kapacitou. Ideálně by měly být zahrnuty také a intermodální propojení tak, aby bylo možné modelovat přestupy ve veřejné dopravě a propojení mezi různými dopravními módy. Přidělování na dopravní síť je dynamický proces, v rámci kterého se optimalizuje rozdělení dopravy tak, aby splňovalo kritérium Nashova ekvilibria (de Dios Ortúzar & Willumsen, 2011). Ve skutečnosti i volbu konkrétní cesty má smysl zkoumat z pohledu diskrétních voleb a pomoci tak nastavit parametry dopravní infrastruktury, které o volbě cesty rozhodují. Ví se, že kromě délky cesty a cestovního času hraje roli vnímaná bezpečnost cesty (Rizzi & de Dios Ortúzar, 2003), subjektivní kvalita infrastruktury a prostředí (Tilahun et al., 2007), aktuální informace, které má aktér k dispozici (Abdel-Aty, Kitamura & Jovanis, 1997) nebo kognitivní chyby, kterých se při volbě trasy cestující dopouští (Šimeček, 2014).

Model dopravní poptávky modeluje kompletní dopravní chování obyvatel modelované oblasti. Zahrnuje model generování cest, distribuce cest a volbu dopravního proudu. Generování cest je řešeno obvykle tak, že se průměrné počty cest a jejich účely vynásobí počtem obyvatel v daném území, popřípadě se tyto údaje nějak spojí s jejich sociodemografickou strukturou. Cílem dopravního modelu totiž obvykle není sledovat dopravní chování jednotlivce, ale spíše zkoumat agregované dopravní proudy. Přesto se rozvíjí také takzvané disagregované dopravní modely, ve kterých dává smysl modelovat vznik poptávky po dopravním chování na úrovni jednotlivce. K tomu se někdy používají rozhodovací modely založené na markovských řetězcích, kdy pravděpodobnost výběru nějaké alternativy závisí na vybraných alternativách v minulosti (Choudhury et al., 2010; Šimeček, 2013). Modely s markovskými řetězci jsou řešeny i v rámci modelů diskrétních voleb (Ben-Akiva & Lerman, 1985), kterými mohou být data o plánování dopravních aktivit jednotlivce modelována. Model distribuce cest je obvykle v dopravních modelech založen na takzvaných gravitačních modelech (Wilson, 1969, 1971, 1998). Ty jsou založeny na tom, že jednotlivé zóny mají určitou přitažlivou sílu k dopravnímu chování za daným účelem (tzv. atraktivitu). Tak například zóna, ve které se nachází velký supermarket, bude mít velkou atraktivitu pro nakupování. Pravděpodobnost, že určitý člověk tuto zónu navštíví za účelem nákupu, závisí dále zhruba řečeno na jeho vzdálenosti od této zóny ale také na ostatních zónách v okolí (jejich atraktivitách a vzdálenostech). S nimi naše zóna v přitažlivosti soutěží. Tento model lze rovněž chápat jako model diskrétních voleb. Člověk si při takovém výkladu volí, do které zóny pojedete nakupovat. Ben-Akiva s kolektivem proto vytvořili disagregovaný na volbě založený model distribuce cest (Ben-Akiva, Gunn & Silman, 1984). Modely volby dopravního módu v dopravních modelech vypadají podobně, jako je na několika místech ukázáno v této knize (např. kapitola 7), liší se pouze v parametrech, které používají. V praxi se modely dopravní poptávky obvykle odhadují na základě dat z cestovních deníků. Jde tedy o data projevených preferencí (*RP*), pro které je nutné modelovat vlastnosti všech alternativ. Na rozdíl od modelů, které vznikají na základě průzkumů vyjádřených preferencí, pracují pouze s tím, co je možné o území a jeho obyvatelích zjistit.

### 7.3 Další studium

K dispozici je celá řada literatury k dalšímu studiu, pochopitelně převážně v angličtině. Patrně základní knihou o analýze diskrétních voleb je (Ben-Akiva & Lerman, 1985). Jde o dosti náročné čtení, zaměřené především na analýzu dat. Je zde popsána celá řada modelů a přístupů, ale vznikají samozřejmě další.

Praktická kniha je od Henschera et al. (2007). Zde je vše od tvorby experimentálního designu po základní analýzy. Kombinované *SP/RP* modely jsou popsány například v (Henscher et al., 2007) kapitole 14. Různé příklady kombinací zdrojů dat jsou v práci Ben-Akivy a kolektivu (1994). Modely s náhodnou proměnnou, tedy *mixed logit* modely, jsou popsány také v Henscher & Greene (2001).

Jak bylo řečeno v úvodu, heuristiky samy o sobě nejsou obvykle předmětem aplikovaného výzkumu, který využívá analýzu rozhodování, je však s jejich působením nutné počítat. Velmi podrobně se vlivem heuristik na *SP* experimenty zabývají Leong a Hensher (2012a, 2012b).

Vyplatí se pročitat také praktické studie, které používají *SP* experiment a analýzu rozhodování. V nich lze kromě osvojení standardního postupu také občas narazit na zajímavé zvláštnosti.

## Literatura

- Aarts, H. A. G., Verplanken, B., & Knippenberg, van, A. F. M. (1998). Predicting behavior from actions in the past: repeated decision making or a matter of habit? *Journal of Applied Social Psychology*, 28(15), 1355-1374. DOI: 10.1111/j.1559-1816.1998.tb01681.x
- Abdel-Aty, M. A., Kitamura, R., & Jovanis, P. P. (1997). Using stated preference data for studying the effect of advanced traffic information on drivers' route choice. *Transportation Research Part C: Emerging Technologies*, 5(1), 39-50. DOI: 10.1016/S0968-090X(96)00023-X
- Adamowicz, W., Swait, J., Boxall, P., Louviere, J., & Williams, M. (1997). Perceptions versus objective measures of environmental quality in combined revealed and stated preference models of environmental valuation. *Journal of environmental economics and management*, 32(1), 65-84. DOI: 10.1006/jeem.1996.0957
- Alberini, A., Ščasný, M. (2011). Context and the VSL: Evidence from a Stated Preference Study in Italy and the Czech Republic, *Environ Resource Econ*, 49, 511–538. DOI: 10.1007/s10640-010-9444-8
- Athira, I. C., Muneera, C. P., Krishnamurthy, K., Anjaneyulu, M. V. L. R. (2016). Estimation of value of travel time for work trips. *Transportation Research Procedia*, 17, 116-123. DOI: 10.1016/j.trpro.2016.11.067
- Axhausen, K. W., & Polak, J. W. (1991). Choice of parking: Stated preference approach. *Transportation*, 18 (1), 59–81. DOI: 10.1007/BF00150559
- de Bekker-Grob EW, et al. (2015). Protective Behaviour of Citizens to Transport Accidents Involving Hazardous Materials: A Discrete Choice Experiment Applied to Populated Areas nearby Waterways. *PLoS ONE* 10(11): e0142507. DOI: 10.1371/journal.pone.0142507
- de Blaeij, A., Florax, R. J., Rietveld, P., Verhoef, E. (2003). The value of statistical life in road safety: A meta-analysis. *Accident Analysis and Prevention* 35 (6) 973–986. DOI: 10.1016/S0001-4575(02)00105-7
- Ben-Akiva, M. E., & Lerman, S. R. (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand*. MIT Press.
- Ben-Akiva, M., & Bierlaire, M. (1999). Discrete choice methods and their applications to short term travel decisions. In *Handbook of transportation science* (pp. 5-33). Springer, Boston, MA.
- Ben-Akiva, M., Bradley, M., Morikawa, T., Benjamin, J., Oppewal, H., Novak, T., & Rao, V. (1994). Combining Revealed and Stated Preferences Data. *Marketing Letters*, 5(4), 335–350.
- Ben-Akiva, M., Gunn, H. F., & Silman, L. a. (1984). Disaggregate trip distribution models. *Doboku Gakkai Ronbunshu*, 1(347), 1–17. DOI: 10.2208/jscej.1984.347\_1
- Beusterien, K. M., Dziekan, K., Flood, E., Harding, G., & Jordan, J. C. (2005). Understanding patient preferences for HIV medications using adaptive conjoint analysis: feasibility assessment. *Value in Health*, 8(4), 453-461. DOI: 10.1111/j.1524-4733.2005.00036.x

- Bierlaire, M. (2003). BIOGEME: A free package for the estimation of discrete choice models. *Proceedings of the 3rd Swiss Transportation Research Conference, Ascona, Switzerland*. OI: 10.1.1.69.4675
- Bierlaire, M. (2016) *PythonBiogeme: a short introduction*. Report TRANSP-OR 160706, Series on Biogeme. Transport and Mobility Laboratory, School of Architecture, Civil and Environmental Engineering, Ecole Polytechnique Fédérale de Lausanne, Switzerland.
- Bischoff, J., & Maciejewski, M. (2016). Simulation of city-wide replacement of private cars with autonomous taxis in Berlin. *Procedia computer science*, 83, 237-244. DOI: 10.1016/j.procs.2016.04.121
- Brownell, C., & Kornhauser, A. (2014). A driverless alternative: fleet size and cost requirements for a statewide autonomous taxi network in New Jersey. *Transportation research record*, 2416(1), 73-81. DOI: 10.3141/2416-09
- Bunch, D. S., Bradley, M., Golob, T. F., Kitamura, R., Occhiuzzo, G. (1993). Demand for Clean-fuel Vehicles in California: A Discrete-Choice Stated Preference Pilot Project. *Transportation Research*, 27A (3), 237–53. DOI: 10.1016/0965-8564(93)90062-P
- Carrion, C., & Levinson, D. (2012). Value of travel time reliability: A review of current evidence. *Transportation research part A: policy and practice*, 46(4), 720-741. DOI: 10.1016/j.tra.2012.01.003
- Carson, R. T., & Louviere, J. J. (2011). A common nomenclature for stated preference elicitation approaches. *Environmental and Resource Economics*, 49(4), 539-559. DOI: 10.1007/s10640-010-9450-x
- CDV & MEDIAN (2017). *Průzkumy poptávky a dopravního chování letištních cestujících a cestujících v příměstské dopravě*.
- CDV & FOCUS (2017). *Sociologický výzkum – Dopravní chování obyvatel města Brna a Brněnské metropolitní oblasti*. [http://iti.brno.cz/wp-content/uploads/2017/10/Dopravni\\_chovani\\_MMB\\_zaverecna\\_zprava.pdf](http://iti.brno.cz/wp-content/uploads/2017/10/Dopravni_chovani_MMB_zaverecna_zprava.pdf)
- Chartrand, T. L. (2005). The role of conscious awareness in consumer behavior. *Journal of Consumer Psychology*, 15(3), 203-210. DOI: 10.1207/s15327663jcp1503\_4
- Choudhury, C. F., Ben-Akiva, M., Abou-Zeid, M. (2010). Dynamic latent plan models. *Journal of Choice Modelling*, 3(2), 50-70. DOI: 10.1016/S1755-5345(13)70035-6
- Croissant, Y. (2011). *Package 'mlogit'*. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.473.2433&rep=rep1&type=pdf>
- Croissant, Y. (2012). Estimation of multinomial logit models in R: The mlogit Packages. *R package version 0.2-2*. <http://cran.r-project.org/web/packages/mlogit/vignettes/mlogit.pdf>
- Danielis, R., Marcucci, E., & Rotaris, L. (2005). Logistics managers' stated preferences for freight service attributes. *Transportation Research Part E: Logistics and Transportation Review*, 41(3), 201-215. DOI: DOI: 10.1016/j.tre.2004.04.003
- de Jong, G., Kroes, E., Plasmeijer, R., Sanders, P., & Warffemius, P. (2004). The Value of Reliability. <http://www.significance.nl/papers/2004-ETC-Value-of-Reliability.pdf>



- de Jong, G., Kouwenhoven, M., Kroes, K., & Rietveld, P. (2009). Preliminary Monetary Values for the Reliability of Travel Times in Freight Transport. *EJTIR* 9(2), 83–99. [www.ejtir.tbm.tudelft.nl](http://www.ejtir.tbm.tudelft.nl)
- de Jong, G., Tseng, Y., Kouwenhoven, M., Verhoef, E., Bates, J. (2007). The Value of Travel Time and Travel Time Reliability Survey Design Final report Prepared for The Netherlands Ministry of Transport, Public Works and Water Management.
- de Dios Ortuzar, J., & Willumsen, L. G. (2011). *Modelling transport*. John Wiley & Sons.
- Dell’Olio, L., Ibeas, A., & Moura, J. L. (2009). Paying for parking: improving stated-preference surveys. *Proceedings of the Institution of Civil Engineers - Transport*, 162(1), 39–45: DOI: 10.1680/tran.2009.162.1.39
- Dubin, J. A., & McFadden, D. L. (1984). An econometric analysis of residential electric appliance holdings and consumption. *Econometrica: Journal of the Econometric Society*, 345–362. DOI: 10.2307/1911493
- Ewing, G., & Sarigöllü, E. (2000). Assessing consumer preferences for clean-fuel vehicles: A discrete choice experiment. *Journal of public policy & marketing*, 19(1), 106–118. DOI: 10.1509/jppm.19.1.106.16946
- Fowkes, A. S. (2007). The design and interpretation of freight stated preference experiments seeking to elicit behavioral valuations of journey attributes. *Transportation Research Part B: Methodological*, 41(9), 966–980. DOI: 10.1016/j.trb.2007.04.004
- Fowkes, A. S., Firmin, P. E., Tweddle, G., & Whiteing, A. E. (2004). How highly does the freight transport industry value journey time reliability—and for what reasons? *International Journal of Logistics: Research and Applications*, 7(1), 33–43. DOI: 10.1080/13675560310001619259
- Fowkes, A.S., Shinghal, N. (2002). The Leeds Adaptive Stated Preference Methodology. *Bilingualism*, 110, 115–122.
- Fraenkel, L., Bodardus, S., & Wittink, D. R. (2001). Understanding patient preferences for the treatment of lupus nephritis with adaptive conjoint analysis. *Medical care*, 39(11), 1203–1216. DOI: DOI: 10.1097/00005650-200111000-00007
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological review*, 62 (1), 650–669. DOI: 10.1093/acprof:oso/9780199744282.003.0002
- Green P, Rao V (1971) Conjoint measurement for quantifying judgmental data. *J Mark Res* 8(3): 355–363. DOI: 10.1177/002224377100800312
- Greene, W. H. (2002). *Nlogit*. <http://pages.stern.nyu.edu/~wgreene/DiscreteChoice/Software/NLOGIT5Referencel.pdf>
- Hanemann, W. M. (1991). Willingness to pay and willingness to accept: how much can they differ? *The American Economic Review*, 81(3), 635–647. DOI: 10.1257/000282803321455449
- Hensher, D. (2001). Measurement of the Valuation of Travel Time Savings, *J. Transp. Econ. Policy*. 35 (2001) 71–98.



- Hensher, D., & Greene, W. (2001). *The Mixed Logit Model: The State of Practice and Warnings for the Unwary*, (October 2011), 39.
- Hensher, D. A., Li, Z., & Rose, J. M. (2013). Accommodating risk in the valuation of expected travel time savings. *Journal of advanced transportation*, 47(2), 206-224. DOI: 10.1002/atr.160
- Hensher, D., Louviere, J., & Swait, J. (1999). Combining sources of preference data. *Journal of Econometrics* (Vol. 89).
- Henscher, D. A., Rose, J. M., & Greene, W. H. (2007). *Applied Choice Analysis: A Primer* (Vol. 102). Cambridge, UK: Cambridge: University Press.
- Hess, S., Bierlaire, M., & Polak, J. W. (2005). Estimation of value of travel-time savings using mixed logit models. *Transportation Research Part A: Policy and Practice*, 39(2-3), 221-236. DOI: 10.1016/j.tra.2004.09.007
- Hess, S., Rose, J. M., & Hensher, D. A. (2008). Asymmetric preference formation in willingness to pay estimates in discrete choice models. *Transportation Research Part E: Logistics and Transportation Review*, 44(5), 847-863. DOI: 10.1016/j.tre.2007.06.002
- Hole, A. R. (2007). Estimating mixed logit models using maximum simulated likelihood. *Stata Journal*, 7(3), 388-401. [https://www.sheffield.ac.uk/polopoly\\_fs/1.105585!/file/mixlogit.pdf](https://www.sheffield.ac.uk/polopoly_fs/1.105585!/file/mixlogit.pdf)
- Iragüen, P., & de Dios Ortúzar, J. (2004). Willingness-to-pay for reducing fatal accident risk in urban areas: an Internet-based Web page stated preference survey. *Accident Analysis & Prevention*, 36(4), 513-524. DOI: 10.1016/S0001-4575(03)00057-5
- Jamson, S., Wardman, M., Batley, R., & Carsten, O. (2008). Developing a driving Safety Index using a Delphi stated preference experiment. *Accident Analysis & Prevention*, 40(2), 435-442. DOI: 10.1016/j.aap.2007.07.014
- Jara-Díaz, S. R. (1998). Time and income in travel demand: towards a microeconomic activity framework. *Theoretical Foundations of Travel Choice Modelling*. Elsevier. <https://www.cec.uchile.cl/~dicedet/sjara/Time%20and%20Income.PDF>
- Jaspers Team (2014). *JASPERS Appraisal Guidance (Transport) The Use of Transport Models in Transport Planning and Project Appraisal*. [http://kc-sump.eu/wordpress/wp-content/uploads/2015/04/Uпотреба-Modela-u-prometnom-planiranju\\_JASPERS\\_kolovoz-2014.pdf](http://kc-sump.eu/wordpress/wp-content/uploads/2015/04/Uпотреба-Modela-u-prometnom-planiranju_JASPERS_kolovoz-2014.pdf)
- Johnston, R. J., Boyle, K. J., Adamowicz, W., Bennett, J., Brouwer, R., Cameron, T. A., & Tourangeau, R. (2017). Contemporary guidance for stated preference studies. *Journal of the Association of Environmental and Resource Economists*, 4(2), 319-405. DOI: doi.org/10.1086/691697
- Kahneman, D., Tversky A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica: Journal of the Econometric Society* 47(3) 263-291. DOI: 10.2307/1914185
- Kouřil, P., Šenk, P., Biler, S. (2014). *Metodika aktivně-cestovního průzkumu*. Brno: Centrum dopravního výzkumu, v. v. i., 48 s. Zadavatel: Technologická agentura ČR.

- Koppelman, F. S., & Bhat, C. (2006). A self instructing course in mode choice modeling: multinomial and nested logit models. [http://caee.utexas.edu/prof/bhat/COURSES/LM\\_Draft\\_060131Final-060630.pdf](http://caee.utexas.edu/prof/bhat/COURSES/LM_Draft_060131Final-060630.pdf)
- Lenk, P. (2011). A comparison of generalized multinomial logit (GMNL) and latent class approaches to studying consumer heterogeneity with some extensions of the GMNL model by Peter J. Lenk. *Applied Stochastic Models in Business and Industry*, 27(6), 580-583. DOI: 10.1002/asmb.941
- Leong, W., & Hensher, D. A. (2012a). Embedding decision heuristics in discrete choice models: A review. *Transport Reviews*, 32(3), 313-331. DOI: 10.1080/01441647.2012.671195
- Leong, W., & Hensher, D. A. (2012b). Embedding multiple heuristics into choice models: An exploratory analysis. *Journal of choice modelling*, 5(3), 131-144. DOI: 10.1016/j.jocm.2013.03.001
- Liew, N., & Chu, J. (2016). How much is your daily commute? Developing a working model to estimate the total travel cost. In *Australasian Transport Research Forum (ATRF), 38th, 2016, Melbourne, Victoria, Australia*. [https://atrf.info/papers/2016/files/ATRF2016\\_Full\\_papers\\_resubmission\\_68.pdf](https://atrf.info/papers/2016/files/ATRF2016_Full_papers_resubmission_68.pdf)
- Loo, B. P. (2002). Role of stated preference methods in planning for sustainable urban transportation: State of practice and future prospects. *Journal of Urban Planning and Development*, 128(4), 210-224. DOI: 10.1061/(ASCE)0733-9488(2002)128:4(210)
- Louviere, J. J., Hensher, D. A., & Swait, J. D. (2000a). *Combining sources of preference data*. *Journal of Econometrics*, 89 (1-2). 197-221.
- Louviere, J., Hensher, D. A., & Swait, J. D. (2000b). *Stated choice methods: analysis and applications*. Cambridge, UK: Cambridge University Press.
- Máca, V., Melichar, J., Urban, J., Braun Kohlová, M., Ščasný, M., Škopková, H., Moldan, B., Jedlička, J., Křivánek, V., Dostál, I., Dufek, J., Hamanová, J., Mach, Š. (2012). *Kvantifikace externích nákladů dopravy v podmínkách České republiky, Závěrečná zpráva projektu CG712-111-520, Praha: Univerzita Karlova*. [https://czp.cuni.cz/czp/images/stories/Vystupy/TranExt/ZZ\\_TrانExt\\_final.pdf](https://czp.cuni.cz/czp/images/stories/Vystupy/TranExt/ZZ_TrانExt_final.pdf)
- Manski, C. F. (1977). The structure of random utility models. *Theory and Decision* 8 (3): 229-254. DOI: 10.1007/BF00133443
- Manski, C. F. (1981). Structural models for discrete data: The analysis of discrete choice. *Sociological methodology*, 12, 58-109.
- Martin, N., & Morich, K. (2011). Unconscious mental processes in consumer choice: Toward a new model of consumer behavior. *Journal of Brand Management*, 18(7), 483-505. DOI: 10.1057/bm.2011.10
- Martin, W. A. & McGuckin, N. A. (1998). *Travel estimation techniques for urban planning* (No. 365). Washington, DC: National Academy Press.
- McFadden, D. L. (1984). Econometric analysis of qualitative response models. *Handbook of econometrics*, 2, 1395-1457. DOI: 10.1016/S1573-4412(84)02016-X
- McFadden, D. (1986). The choice theory approach to market research. *Marketing science*, 5(4), 275-297. DOI: 10.1287/mksc.5.4.275

- Merino-Castello, A. (2003). Eliciting consumers preferences using stated preference discrete choice models: contingent ranking versus choice experiment. *UPF Economics and Business Working Paper No. 705*. DOI: 10.2139/ssrn.562982
- Metz, D. (2008). The myth of travel time saving. *Transport reviews*, 28(3), 321-336. DOI: 10.1080/01441640701642348
- Mrozek, J. R., Taylor, L. O. (2002). What determines the value of life? A meta-analysis. *Journal of Policy Analysis and Management*, 21 (2) 253–270. DOI: 10.1002/pam.10026
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*, Vienna, Austria, 2018. <https://www.R-project.org/>.
- Ramjerdi, F., Rand, L., Sætermo, I. A. F., & Sælensminde, K. (1997). The Norwegian value of time study, part I. *Institute of Transport Economics, Oslo*.
- Richardson, A. (2002). Simulation study of estimation of individual specific values of time by using adaptive stated-preference survey. *Transportation Research Record: Journal of the Transportation Research Board*, (1804), 117-125. DOI: 10.3141/1804-16
- Rizzi, L. I., & de Dios Ortúzar, J. (2003). Stated preference in the valuation of interurban road safety. *Accident Analysis & Prevention*, 35(1), 9-22. DOI: 10.1016/S0001-4575(01)00082-3
- Sarrias, M. (2016). Discrete choice models with random parameters in R: The Rchoice Package. *Journal of Statistical Software*, 74(10), 1-31. DOI: 10.18637/jss.v074.i10
- Sarrias, M., & Daziano, R. A. (2017). Multinomial logit models with continuous and discrete individual heterogeneity in R: the gnm1 package. *Journal of Statistical Software*, 79(2), 1-46. DOI: 10.18637/jss.v079.i02
- Shams, K. (2016). Understanding the Value of Travel Time Reliability for Freight Transportation to Support Freight Planning. *FIU Electronic Theses and Dissertations*. 2828. DOI: 10.25148/etd.FIDC001255
- Shogren, J. F., Shin, S. Y., Hayes, D. J., & Kliebenstein, J. B. (1994). Resolving differences in willingness to pay and willingness to accept. *The American Economic Review*, 255-270.
- Small, K. A. (2012). Valuation of Travel Time, *Econ. Transp.* 4, 2012. doi:10.1016/j.ecotra.2012.09.002.
- Soukup, P. (2013). Věcná významnost výsledků a její možnosti měření. *Data a výzkum-SDA Info*, 7(2), 125-148. DOI: 10.13060/23362391.2013.127.2.41
- Star EU & CDV (2014). Štúdia realizovateľnosti – aktualizácia, úsek Žilina – Košice – Čierna nad Tisou št. hr.
- Šikl, R., & Šimeček, M. (2015). Visual space perception at different levels of depth description. *Attention, Perception, & Psychophysics*, Volume 77, Issue 6, pp 2098-2107. DOI: 10.3758/s13414-015-0917-2
- Šimeček, M. (2013) Dopravní chování jako úkol pro agentní simulaci. In Kelemen, J., Rybár, J., Farkaš, I., Takáč, M. [eds.] *Kognitivní věda a umělý život*, FPF Slezská univerzita, ISBN 978-80-7248-863-6, s. 249-253.

- Šimeček, M. (2014). Human path assignation at maps: discrete choice model of visual attributes of alternatives. *In the Proceedings of 10th International Conference on Travel Survey Methods*.
- Šimeček, M., & Dufek, J. (2016). A Freight Modal Shift Model for Slovakia. *Transportation Research Procedia*, 14, 2814-2819, DOI: 10.1016/j.trpro.2016.05.347
- Tanner, R. J., Ferraro, R., Chartrand, T. L., Bettman, J. R., & Baaren, R. V. (2007). Of chameleons and consumption: The impact of mimicry on choice and preferences. *Journal of Consumer Research*, 34(6), 754-766. DOI: <http://dx.doi.org/10.1086/522322>
- Thurstone, L. L. (1994). A law of comparative judgment. *Psychological Review*, 101(2), 266. <http://mlab.no/blog/wp-content/uploads/2009/07/thurstone94law.pdf>
- Tichá, Š. (2004). *Strojírenská metrologie část 1. Ostrava: VŠB-Technická univerzita*. <http://books.fs.vsb.cz/StrojMetro/strojirenska-metrologie.pdf>
- Tilahun, N. Y., & Levinson, D. M. (2010). A Moment of Time: Reliability in Route Choice Using Stated Preference. *Journal of Intelligent Transportation Systems*, 14(3), 179-187. DOI: 10.1080/15472450.2010.484751
- Tilahun, N. Y., Levinson, D. M., & Krizek, K. J. (2007). Trails, lanes, or traffic: Valuing bicycle facilities with an adaptive stated preference survey. *Transportation Research Part A: Policy and Practice*, 41(4), 287-301. DOI: 10.1016/j.tra.2006.09.007
- Train, K. (2001). A comparison of hierarchical bayes and maximum simulated likelihood for mixed logit. *Working Paper*, Department of Economics, University of California, Berkeley. <https://emlab.berkeley.edu/~train/compare.pdf>
- Train, K. E., McFadden, D. L., & Ben-Akiva, M. (1987). The demand for local telephone service: A fully discrete model of residential calling patterns and service choices. *The RAND Journal of Economics*, 109-123. DOI: 10.2307/2555538
- Train, K., & Croissant, Y. (2012). Kenneth Train's exercises using the mlogit package for R. R, 25, 0-2. <http://meetings-archive.debian.net/mirror/CRAN/web/packages/mlogit/vignettes/Exercises.pdf>
- Truong, T. P., & Hensher, D. A. (1985). Measurement of travel time values and opportunity cost from a discrete-choice model. *The Economic Journal*, 438-451. DOI: 10.2307/2233219
- Tseng, Y. Y., Yue, W. L., & Taylor, M. A. (2005). The role of transportation in logistics chain. *Eastern Asia Society for Transportation Studies*, 5, 1657-1672. <http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=A78C9D44C4CFFBD91B452338EBC20C14?doi=10.1.1.184.8769&rep=rep1&type=pdf>
- Viscusi, W. K., Aldy, J. E. (2003). The value of a statistical life: A critical review of market estimates throughout the world, *Journal of Risk and Uncertainty* 27 (1) 5-76. DOI: 10.3386/w9487
- Viton, P. A. (2014). *Discrete-Choice Logit Models with R*. <http://www.r-project.org/>
- Von Neumann, J., & Morgenstern, O. (2007). *Theory of games and economic behavior (commemorative edition)*. Princeton university press. <https://pdfs.semanticscholar.org/0375/379194a6f34b818962ea947bff153adf621c.pdf>

- Vorel, J., & Maier, K. (2007). *Learning the public preferences for living environment characteristics: the experimental approach*.
- Wardman, M., Chintakayala, P., de Jong, G., & Ferrer, D. (2012). *European wide meta-analysis of values of travel time*.
- Wheeler, B., & Braun, M. J. (2015). *Package 'AlgDesign'*. <https://mirror.its.sfu.ca/mirror/CRAN/web/packages/AlgDesign/AlgDesign.pdf>
- Widmer, P. (2004). *Einfluss von Aenderungen des Parkierungs-Angebotes auf das Verkehrsverhalten*. Ecole Polytechnique Fédérale de Lausanne (EPFL).
- Wilson, A. G. (1969). The Use of Entropy Maximising Models, in the Theory of Trip Distribution, Mode Split and Route Split. *Journal of Transport Economics and Policy*, 3, 108-126. DOI: 10.2307/20052128
- Wilson, A. G. (1971). Family of spatial interaction models, and associated developments. *Environment and Planning*, 3(1), 1–32. DOI: 10.1068/a030001
- Wilson, A. G. (1998). Land-use/Transport Interaction Models: Past and Future. *Journal of Transport Economics and Policy*, 32(1).

