

UNIVERZITA KARLOVA V PRAZE

FAKULTA SOCIÁLNÍCH VĚD

Institut sociologických studií

Anna Krčková

**Bayesovská statistika – limity a možnosti
využití v sociologii**

Diplomová práce

Praha 2014

Autor práce: **Anna Krčková**

Vedoucí práce: **PhDr. Mgr. Ing. Petr Soukup**

Rok obhajoby: 2014

Bibliografický záznam

KRČKOVÁ, Anna. *Bayesovská statistika – limity a možnosti využití v sociologii*. Praha, 2014. 73 s. Diplomová práce (Mgr.) Univerzita Karlova, Fakulta sociálních věd, Institut sociologických studií. Katedra sociologie. Vedoucí diplomové práce PhDr. Mgr. Ing. Petr Soukup.

Abstrakt

Cílem této diplomové práce je zjistit, jak lze využít bayesovskou statistiku v analýze sociologických dat a porovnat výsledky frekvenčního a bayesovského přístupu. Bayesovská statistika vztahuje na statistické parametry pravděpodobnostní rozdělení. V bayesovském přístupu je na začátku analýzy parametrům přiřazena tzv. apriorní pravděpodobnost volená na základě všech známých relevantních informací. Po zkombinování apriorní pravděpodobnosti a napozorovaných dat se pak počítá tzv. posteriorní pravděpodobnost, pomocí které vytváříme statistické úsudky. Komparace zmíněných přístupů je provedena jak z hlediska teoretických východisek a jejich postupů, tak pomocí analýz na konkrétních sociologických datech. Porovnány jsou bodové a intervalové odhady, testování hypotéz (na příkladu dvouvýběrového t-testu) a vícenásobná lineární regresní analýza. Výsledkem je zjištění, že vzhledem k filozofii a díky interpretační jednoduchosti je bayesovská analýza pro sociologická data vhodnější než frekvenční analýza. Porovnání analýz ukázalo, že výsledky frekvenční a objektivní bayesovské analýzy se od sebe neliší bez ohledu na velikost výběru. Pro testování hypotéz můžeme v bayesovské statistice použít kredibilní intervaly. Použití subjektivní bayesovské analýzy u menších datových souborů má na výsledky větší vliv, a naopak u větších datových souborů je apriorní informace převážena daty. Zároveň jsou výsledky subjektivní bayesovské analýzy přesnější.

Abstract

The purpose of this thesis is to find how we can use Bayesian statistics in analysis of sociological data and to compare outcomes of frequentist and Bayesian approach. Bayesian statistics uses probability distributions on statistical parameters. In the beginning of the analysis in Bayesian approach a prior probability (that is chosen on the basis of relevant information) is attached to the parameters. After combining prior probability and our observed data, posterior probability is computed. Because of the posterior probability we can make statistical conclusions. Comparison of approaches was made from the view of theoretical foundations and procedures and also by means of analysis of sociological data. Point estimates, interval estimates, hypothesis testing (on the example of two-sample t-test) and multiple linear regression analysis were compared. The outcome of this thesis is that considering its philosophy and thanks to the interpretational simplicity the Bayesian analysis is more suitable for sociological data analysis than common frequentist approach. Comparison showed that there is no difference between outcomes of frequentist and objective Bayesian analysis regardless of the sample size. For hypothesis testing we can use Bayesian credible intervals. Using subjective Bayesian analysis on smaller samples has bigger impact on the outcomes and vice versa and it also offers more specific outcomes.

Klíčová slova

Bayesovská statistika, Bayesova věta, komparace, statistika, frekvenční přístup

Keywords

Bayesian statistics, Bayesian theorem, comparison, statistics, frequentist approach

Rozsah práce: 114 684 znaků

Prohlášení

1. Prohlašuji, že jsem předkládanou práci zpracovala samostatně a použila jen uvedené prameny a literaturu.
2. Prohlašuji, že práce nebyla využita k získání jiného titulu.
3. Souhlasím s tím, aby práce byla zpřístupněna pro studijní a výzkumné účely.

V Praze dne 5. 5. 2014

Anna Krčková

Poděkování

Na tomto místě bych ráda poděkovala PhDr. Mgr. Ing. Petru Soukupovi za cenné připomínky při vedení práce a svým blízkým za podporu při studiu.

Projekt diplomové práce

Bc. Anna Krčková

Fakulta sociálních věd

Sociologie – Aplikovaný sociologický výzkum a jeho metodologie

1. Předpokládaný název práce.

Bayesovská statistika – limity a možnosti využití v sociologii

2. Námět práce zahrnující formulaci a vstupní diskusi poznávacího problému.

Diplomová práce se bude zabývat možnostmi využití bayesovského přístupu v analýze sociálně-vědních kvantitativních dat. Bayesovská statistika, v jejímž jádru je Bayesova věta, stojí v opozici proti v současnosti převládajícímu frekvenčnímu (klasickému) přístupu v analýze dat a statistickému uvažování v sociologii. Při testování hypotéz ve frekvenčním přístupu se nejdříve stanoví hypotézy – pokud se tyto hypotézy shodují s daty, lze je prohlásit za platné, pokud odlišnost přesáhne určitou stanovenou velikost, jsou prohlášeny za neplatné. V bayesovském přístupu je na začátku každé teorii přiřazena subjektivní pravděpodobnost její platnosti – tzv. apriorní pravděpodobnost (volená na základě všech známých relevantních informací), při porovnávání dat s teorií se pak zpřesňují výchozí hypotézy podle dat a počítá se podmíněná, tzv. aposteriorní pravděpodobnost. Bayesiáni také hovoří o potvrzených nebo vyvrácených teoriích, avšak na základě pravděpodobnosti velmi velké (blížící se jedné) či velmi malé (blížící se nule). Bayesovský přístup se využívá v mnoha oborech (ve financích, lékařství, kriminalistice ad.), jeho využití je možné i v sociologii – jedním z cílů této práce bude představit konkrétní metody, které jsou pro analýzu sociálně-vědních dat vhodné.

3. Předpokládané metody zpracování a předběžná struktura práce.

Na základě studia odborné literatury na začátku vyložím podrobné teoretické uvedení do Bayesovské statistiky včetně diskuse nad jejími výhodami a kritikou. Dále se budu zabývat porovnáním se současným frekvenčním (klasickým) přístupem. V hlavní části nastíním konkrétní možnosti využití a limity tohoto přístupu v sociologii. Na základě příkladů analýzy konkrétních kvantitativních sociologických dat představím

rozdílnosti obou přístupů a popíše případy, kdy je a kdy není Bayesovský přístup vhodný. V závěrečné diskusi shrnu hlavní zjištění této práce.

4. Orientační seznam literatury.

GILL, Jeff. *Bayesian methods: a social and behavioral sciences approach*. 2nd ed.

Boca Raton: Chapman & Hall CRC, ©2008. xxxvii, 711 s. Statistics in the social and behavioral sciences series. ISBN 1-58488-562-9.

HEBÁK, Petr. Srovnání klasické a bayesovské pravděpodobnosti a statistiky (1.).

In: *ACTA OECONOMICA PRAGENSIA*. Praha: VŠE, 2012, s. 69-87. Dostupné z:

<http://www.vse.cz/aop/abstrakt.php3?IDcl=359>

HOFF, Peter D. *A first course in Bayesian statistical methods* [online]. New York:

Springer, 2009. Springer texts in statistics [cit. 2012-05-28]. Dostupné z:

<<http://www.springerlink.com/content/978-0-387-92299-7>>. ISBN 978-0-387-92407-6.

JACKMANN, Simon. *Bayesian Analysis for the Social Sciences*. Chichester, UK:

Wiley, 2009. ISBN 978-0-470-01154-6.

KRUSCHKE, John K. *Doing Bayesian data analysis: a tutorial with R and BUGS*.

Burlington: Academic Press, ©2011. xvii, 653 s. ISBN 978-0-12-381485-2.

PRESS, S. James a TANUR, Judith M. *The subjectivity of scientists and the bayesian approach*. New York: John Wiley & Sons, 2001. x, 274 s. A Wiley-Interscience

publication. Wiley series in probability and statistics. Texts, references, and pocketbooks section. ISBN 0-471-39685-0.

ROSENTHAL, Jeffrey S. *Zasažen bleskem: podivuhodný svět pravděpodobností*. Vyd.

1. Praha: Academia, 2008. 294 s. Galileo; sv. 20. ISBN 978-80-200-1645-4.

WESTERN, Bruce. Bayesian Analysis for Sociologist: An Introduction. *Sociological Methods & Research*. 1999, 28 (1), 7-34.

5. Jméno konzultanta a jeho písemný souhlas se spoluprací

PhDr. Ing. Petr Soukup

Obsah

| | |
|---|-----------|
| ÚVOD | 4 |
| TEORETICKÁ ČÁST | 7 |
| 1. Proč se zabývat bayesovskou statistikou v sociologii? | 7 |
| 1.1 Vědecký proces v sociálních vědách..... | 7 |
| 2. Statistické usuzování v sociologii | 7 |
| 2.1 Kritika frekvenčního přístupu | 8 |
| 2.2 Proč používáme frekvenční statistiku?..... | 9 |
| 2.3 Představení bayesovského přístupu..... | 10 |
| 2.4 Vhodnost bayesovské statistiky pro sociologická data | 11 |
| 2.5 Omezení | 12 |
| 3. Historie statistiky | 12 |
| 3.1 Počátky statistiky..... | 12 |
| 3.2 Politická aritmetika 16. a 17. století..... | 12 |
| 3.3 Subjektivně orientovaná 2. polovina 18. – 19. století..... | 13 |
| 3.4 Objektivisticky orientovaný konec 19. – 20. století..... | 13 |
| 3.5 První renesance bayesovského přístupu v 2. polovině 20. století | 15 |
| 3.6 Druhá renesance bayesovského přístupu v 90. letech 20. století..... | 16 |
| 3.6.1 MCMC algoritmy a prudký rozvoj | 16 |
| 3.6.2 Výuka statistiky – bayesovsky nebo frekvenčně? | 17 |
| 3.6.3 Situace v ČR | 17 |
| 3.6.4 Budoucí vývoj..... | 18 |
| 4. Základní pojmy a koncepty | 18 |
| 4.1 Pravděpodobnost | 19 |
| 4.1.1 Základní vlastnosti | 19 |
| 4.2 Typy pravděpodobností a jejich využití | 19 |
| 4.2.1 Klasická (Laplaceova) definice pravděpodobnosti..... | 19 |
| 4.2.2 Statistická/frekvenční definice pravděpodobnosti..... | 19 |
| 4.2.3 Subjektivní/bayesovská definice pravděpodobnosti | 20 |
| 4.2.4 Využití jednotlivých přístupů | 22 |
| PRAKTICKÁ ČÁST | 24 |

| | |
|--|-----------|
| 5. Frekvenční vs. bayesovská statistika | 24 |
| 5.1 Frekvenční statistika..... | 24 |
| 5.2 Bayesovská statistika..... | 25 |
| 5.2.1 Interpretativní a analytické základy..... | 26 |
| 5.2.2 Kritika a reakce na ni | 31 |
| 5.2.3 MCMC algoritmy..... | 32 |
| 6. Porovnání analýz | 33 |
| 6.1 Apriorní rozdělení..... | 34 |
| 6.1.1 Pravděpodobnostní rozdělení | 34 |
| 6.1.2 Rozdělení v sociálních vědách | 35 |
| 6.1.3 Rozdělení v bayesovské statistice..... | 36 |
| 6.2 Bodový odhad | 38 |
| 6.2.1 Frekvenční přístup k bodovým odhadům | 38 |
| 6.2.2 Bayesovský přístup k bodovým odhadům | 38 |
| 6.3 Intervalové odhady | 39 |
| 6.3.1 Frekvenční intervaly spolehlivosti | 39 |
| 6.3.2 Bayesovské kredibilní intervaly | 40 |
| 6.3.3 Příklad: Odhady průměrného věku..... | 40 |
| 6.4 Testování hypotéz | 45 |
| 6.4.1 Klasický přístup k testování hypotéz..... | 45 |
| 6.4.2 Bayesovský přístup k testování hypotéz..... | 47 |
| 6.4.3 Příklad: Dvouvýběrový t-test | 48 |
| 6.5 Regresní analýza..... | 52 |
| 6.5.1 Příklad: Lineární regresní analýza | 53 |
| 6.6 Shrnutí..... | 55 |
| 6.7 Software | 56 |
| ZÁVĚR | 57 |
| SUMMARY | 60 |
| POUŽITÁ LITERATURA | 62 |
| SEZNAM OBRÁZKŮ | 66 |
| SEZNAM GRAFŮ | 66 |

| | |
|----------------------------|-----------|
| SEZNAM TABULEK..... | 66 |
| SEZNAM PŘÍLOH | 67 |
| PŘÍLOHY | 68 |

Úvod

Předkládaná diplomová práce se zabývá využitím bayesovské statistiky v sociologii, respektive v analýze sociologických dat. Bayesovská statistika je oblast statistického usuzování, která se v posledních desetiletích v zahraničí prudce rozvíjí, a to nejen v rámci teoretické statistiky, ale i v rámci aplikace na sociálněvědní data (např. Gill, 2008; Jackman, 2009). Na rozdíl od široce rozšířené a běžně používané frekvenční statistiky, není bayesovská statistika v České republice obecně příliš často používaná, v sociologii je pak téměř neznámá.

Dvě hlavní výzkumné otázky této diplomové práce jsou: Jak lze využít bayesovský přístup v analýze sociologických dat? Jak se od sebe liší výsledky frekvenčního a bayesovského přístupu? Hlavními cíly této práce jsou představit bayesovskou statistiku jako alternativní přístup a porovnat jej s frekvenčním přístupem. Bayesovská statistika je pojednána jak z hlediska pravděpodobnostní podstaty a základních konceptů, tak z hlediska statistického postupu. Porovnání je provedeno jak na filozofické a teoretické úrovni, tak na praktické úrovni pomocí příkladů konkrétních analýz. Na základě tohoto porovnání práce odpovídá na výzkumné otázky. Zároveň může tato práce sloužit jako úvod do bayesovské statistiky v sociologii, který zájemce o toto téma provede od základních koncepcí, až po jejich aplikace. Součástí práce je také návod na použití programu BugsXLA, ve kterém je bayesovská analýza v rámci této práce prováděna. Smysl a význam práce tkví v zpracování a představení nového (pro české sociology) druhu statistiky, která je v sociologii dobře využitelná. V tomto smyslu se tato diplomová práce snaží zaplnit „prázdné místo“ v české sociologické literatuře o bayesovské analýze dat.

Práce je strukturována na teoretickou a praktickou část. Teoretická část je rozčleněna na tři tematické oblasti zabývající se postupně smyslem použití bayesovské statistiky na sociologická data, historií statistiky a základy statistické inference. V první tematické oblasti se budeme nejdříve zabývat odpovědí na otázku, proč se bayesovskou statistikou v sociologii vůbec zabývat. Dotkneme se tématu vědeckého procesu v sociálních vědách a postavení statistiky v současné vědě, která dle všeho v dnešní době zažívá svůj „zlatý věk“. Dále popíšeme současnou situaci v analýze sociálněvědních dat, ve které je mechanické používání standardní frekvenční statistiky poměrně ostře kritizováno. Zamyslíme se také nad tím, proč přes uvedené nedostatky frekvenční statistiky u ní nadále setrváváme. Protože je tato práce především o

bayesovské statistice, již v této části alespoň krátce popíšeme její podstatu a postup. Zdůvodníme také, proč jsou sociálněvědní, respektive sociologická, data pro bayesovskou statistiku vhodná. Nakonec v této části uvedeme situace, kdy budou výsledky frekvenční a bayesovské statistiky shodné.

V druhé tematické oblasti o historii statistiky popíšeme vývoj této vědní disciplíny se zaměřením na dvě soupeřící oblasti statistiky – bayesovskou a frekvenční. Zároveň v této části uvedeme, kdy se statistické usuzování stalo nástrojem sociologů.

Ve třetím tematickém celku se budeme zabývat základy statistické inference. Popíšeme základní definice a pravidla pravděpodobnosti a uvedeme definice tří druhů pravděpodobnosti – klasické, frekvenční a subjektivní. Zaměříme se zvláště na frekvenční a subjektivní, které tvoří základ frekvenční a bayesovské statistiky.

V praktické části se již budeme zabývat konkrétním porovnáním frekvenční a bayesovské statistiky. Nejdříve půjde o komparaci teoretickou, která navazuje na poznatky uvedené v teoretické části. Hlavní důraz je kladen na bayesovskou statistiku. Konkrétně na její postup a interpretativní a analytické základy, na kterých je postavena. Část věnovaná podmíněné pravděpodobnosti a Bayesově větě je doplněna o příklady, jejichž cílem je teoretické informace vhodně doplnit. Uvedu také nejčastější kritiku, které bayesiáni čelí, a jejich odpovědi na ni. Popíšu krátce MCMC algoritmy, na nichž jsou složitější bayesovské modely postaveny, a které využívají statistické programy umožňující bayesovské analýzy.

Po krátkém úvodu do pravděpodobnostních rozdělení a bayesovských apriorních rozdělení už bude následovat porovnání přístupů pomocí konkrétních analýz. Frekvenční analýzy jsou prováděny v programu SPSS, bayesovské analýzy jsou prováděny pomocí doplňku do Excelu BugsXLA, který usnadňuje používání programu WinBUGS. Analýzy jsou počítané na datech ze skutečného sociologického výzkumu, konkrétně výzkumu CVVM Naše společnost 2013 - prosinec. Porovnáme bodové a intervalové odhady na základě odhadů průměrného věku respondenta. Na tomto příkladu si ukážeme, jak se liší výsledky bayesovské analýzy za použití různých apriorních rozdělení a zároveň použití těchto rozdělení na různě velké výběry, resp. datové soubory. Následuje porovnání testování hypotéz a ukázka dvouvýběrového t-testu. Aby nezůstalo u zcela základních výpočtů, provedeme také vícenásobnou lineární regresní analýzu, na které opět ilustrujeme různá vstupní apriorní rozdělení a porovnáme výsledky s frekvenčním přístupem. Na konec této části shrneme informace o

srovnávaných přístupech v přehledné tabulce. V závěru ještě uvedeme krátký přehled programů umožňujících bayesovskou analýzu.

Teoretická část

1. Proč se zabývat bayesovskou statistikou v sociologii?

1.1 Vědecký proces v sociálních vědách

Empirická sociologie, a zejména její část zaměřující se na kvantitativní výzkum, se dnes, stejně jako většina dalších empiricky orientovaných oborů, bez statistického usuzování neobejde. Sociologie byla v historii obviňována z toho, že není skutečnou vědou, neboť nepoužívá vědecké metody, které jsou typické pro „skutečné“ přírodní vědy.

Kvantitativní výzkum, respektive výzkumný proces empirické sociologie, používá deduktivní metodu, kterou lze zjednodušeně shrnout do následujícího postupu. Na začátku stojí teoretický nebo praktický problém, který je pomocí výzkumné otázky a konkrétních hypotéz převeden do formy vztahů mezi proměnnými, jež můžeme testovat. Po rozhodnutí o populaci a způsobu výběru se konstruuje konkrétní nástroj pro sběr dat, například dotazník. Následuje sběr dat samotný. Po získání dat přichází na řadu statistika a její metody, které nám umožňují data analyzovat a určit, zda přijmout nebo zamítnout testované hypotézy. Pokud použijeme empiricko-deduktivním model vědecké metody jako kritérium pro určení toho, co je vědecké a co není, pak můžeme snadno klasifikovat velkou část výzkumné aktivity v sociálních vědách jako vědeckou (Gill, 2008, s. 24). Jackman v tomto ohledu pokládá otázku, jestli to nebyla statistika, která pomohla vědeckému statusu sociálních věd (Jackman, 2009, s. xxxiv).

V jádru kvantitativní sociologie tedy stojí empirický výzkum, jehož cílem je získat odpovědi na výzkumné otázky. Není pochyb o tom, že vzhledem k množství posbíraných dat a výzkumů (uložených např. v datových archivech), jsou i sociální vědci aktivně zapojeni v modelu kumulativního poznání¹. Tento kumulativní charakter empirické sociologie je skutečnost, která, jak se v další části této práce dozvíme, nahrává používání bayesovské statistiky v sociologii.

2. Statistické usuzování v sociologii

Statistika je v dnešní době již standardní součástí bakalářských i magisterských studijních programů sociologie na všech univerzitách. Statistika jako taková je však často brána jako nástroj pro analýzu, po jejíchž základech studenti (a možná nejen

¹ Ačkoli např. dle Gilla (2008, s. 25) může být někdy otázkou kvalita a trvalost těchto získaných znalostí.

studenti) často raději nehlobají. To, co si dnešní student sociologie představí pod pojmem statistika je však pouze jedním z přístupů – tzv. frekvenční statistika. Historicky můžeme zjednodušeně vymežit dvě školy statistiky, které spolu dodnes svým způsobem soupeří a mají své zastánce – frekvenční statistiku a bayesovskou statistiku.

2.1 Kritika frekvenčního přístupu

Standardní frekvenční inference předpokládá nekonečné série pokusů, replikací, nebo experimentů používajících stejný *research design*. Její postupy vychází z frekvenční neboli objektivní definice pravděpodobnosti. Tuto definici pravděpodobnosti a její předpoklady, které lze v reálných situacích jen obtížně dodržet, shrnuje Hebák takto: „Při klasickém pojetí neexistuje žádná jiná rozumná možnost, jak pravděpodobnost určit, než provést za stejných podmínek (*je to fyzikálně možné?*) dostatečně velký počet (*nejlépe nekonečný?*) nezávislých (*jak to poznáme, když pojem nezávislosti vyžaduje znalost pravděpodobnosti?*) pokusů a ztotožnit výslednou relativní četnost (*pro jaký počet pokusů je výsledná?*) s neznámou pravděpodobností (*je opravdu zcela neznámá a nic o ní nevíme?*).“ (Hebák, 2012a, s. 76; zvýraznění v originále) Tato perspektiva, která pochází od Neymana a Pearsona, a která byla později zkombinována s Fisherovou fixací *p*-hodnot, je dnes typická pro usuzování v sociálních vědách (Gill, 1999, s. 648).

Velmi málo sociálních vědců, pokud vůbec nějakí, by bylo ochotno vážně tvrdit, že se lidské jednání shoduje s tímto objektivistickým pravděpodobnostním modelem. Myšlenky jako „osobní názor“, „spokojenost“ atd. neexistují nějakým fyzicky hmotným způsobem, který lze v rámci experimentu mnohokrát opakovat. V přímém kontrastu k frekvenční definici je bayesovská či „subjektivistická“ koncepce pravděpodobnosti (více o těchto koncepcích v podkapitole 4.2). Subjektivní pravděpodobnost je chápána jako tzv. stupeň důvěry (*degree of belief*), jež je individuální výzkumník ochoten osobně danému jevu přidělit. Slouží k vyjádření nejistoty o nastání daného jevu, který může být relevantní pouze jedné pozorované situaci nebo experimentu (Gill, 2008, s. 26).

V souvislosti s výše uvedeným, dochází při použití frekvenční statistiky v sociálních vědách k problémům s použitím některých statistických metod a jejich interpretací. Například Hebák (2012a, 2012b, 2012c) například naráží na skutečnost, že „... mezi autory článků v nestatisticky orientovaných časopisech existuje až fetišizovaná představa, že odborný příspěvek neobstojí bez statistického *testu významnosti*.“ (Hebák,

2012a, s. 69; zvýraznění v originále) Gigerenzer (2004) nazývá problém testování hypotéz „rituál nulové hypotézy“ („*null ritual*“). Rituálem jej nazývá, neboť podle něj vyvolává iluzi vědeckosti a jako každý rituál v nás vytváří pocit sounáležitosti se sociální skupinou. Zároveň však způsobuje v sociálních vědách vytrácení statistického uvažování. Tento rituál podle něj spočítává ve třech krocích: stanovení nulové hypotézy jako „žádný rozdíl mezi průměry“ nebo „žádná korelace“, použití 5 % hladiny významnosti a neustálého používání tohoto postupu bez jakéhokoli přemýšlení. Podle Hebáka „... je proto třeba varovat nebo aspoň připomenout odborné veřejnosti skutečnost, že nejen *mechanické používání* nástrojů a terminologie z oblasti statistických úsudků, ale **samotné tyto statistické nástroje** jsou oprávněně odbornou veřejností z různých hledisek už *dlouhodobě* kritizovány a zpochybňovány.“ (Hebák, 2012a, s. 69; zvýraznění v originále)

2.2 Proč používáme frekvenční statistiku?

Důvodů, proč sociologové přetrvávají u frekvenční statistiky, je hned několik. V kontextu České republiky je nejspíš hlavním důvodem neznalost bayesovského přístupu. V zahraničí naopak najdeme poměrně hodně zastánců bayesovské statistiky, přesto mezi sociology ani v zahraničí nedochází k nějakému masovému přijetí tohoto přístupu. Na rozdíl od profesionálních statistiků-metodologů, v kvantitativní sociologii bayesovská statistika nevzkvétá. Western (1999) píše, že to lze částečně vysvětlit různými prioritami metodologů a aplikovaných výzkumníků. Zatímco pro metodology je teoretické pozadí tohoto přístupu zajímavé, empiričtí sociologové se o ně příliš nezajímají (Western, 1999, s. 7-8). Gill (2008) předkládá několik důvodů, proč podle něj výzkumníci v sociálních a behaviorálních vědách stále přetrvávají u tradičního pojetí statistiky:

- parametry populace jejich zájmu jsou skutečně neměnné za všech realistických podmínek,
- nemají jakoukoli předchozí znalost či informaci, kterou by vložili do procesu specifikace modelu,
- je pro ně důležité poskytovat statistické výsledky, jako kdyby data byla pořízena v kontrolovaném experimentu,
- starají se spíše o „významnost“ („signifikanci“) než o velikost účinku (*effect size*),
- upřednostňují velmi automatické procedury typu „kuchařky“ (Gill, 2008, s. 3).

Kromě uvedených důvodů k dominanci frekvenčního pojetí statistiky v sociálních vědách nahrává také skutečnost, že během vývoje moderní statistiky její klíčové postavy měly silné předsudky vůči bayesovskému usuzování (ke kterým však v některých případech vedly spíše důvody osobního rázu²). Obecně můžeme souhlasit s tím, že v mnoha disciplínách používajících frekvenční statistiku je přítomný nedostatek metodologické introspekce (Gill, 2008, s. 3-4).

2.3 Představení bayesovského přístupu

Doposud jsme o bayesovském přístupu³ mluvili pouze jako o alternativě frekvenčního přístupu. Vzhledem k tomu, že tento přístup není příliš známý, je důležité již nyní předestřít jeho filozofii a postup. Bayesovská statistika vztahuje na statistické parametry pravděpodobnostní rozdělení. V bayesovském přístupu je na začátku analýzy každé teorii přiřazena subjektivní pravděpodobnost její platnosti – tzv. apriorní pravděpodobnost (volená na základě všech známých relevantních informací). Toto pravděpodobnostní rozdělení kvantifikuje výzkumníkovu nejistotu o hodnotě parametrů a je přiřazováno předtím, než napozorujeme data. Na základě zkombinování apriorní pravděpodobnosti a napozorovaných dat se pak počítá podmíněná, tzv. posteriorní pravděpodobnost (Rosentah, 2008, s. 2-3). Na základě posteriorní pravděpodobnosti pak vytváříme statistické úsudky. Přirozeně může být apriorní pravděpodobnost s objevením nových informací změněna a proces znovu opakován (Gill, 2008, s. 26). Proces opakování analýzy a zpřesňování výchozí pravděpodobnosti na základě nových informací je základem bayesovské statistiky, který je v souladu s vědeckým procesem jako takovým (Lynch, 2007, s. 50). Kromě toho, že je v souladu s vědeckým procesem, je v souladu také s tím, jak uvažujeme v reálném životě⁴.

Subjektivita apriorního rozdělení je nekritizovanějším bodem bayesovské statistiky. Klíčovou obranou bayesovského přístupu je nepopiratelný fakt, že všechny statistické modely jsou subjektivní. Rozhodnutí například o specifikaci proměnné a

² Více v části 3 o historii statistiky a R. A. Fisherovi.

³ Bayesovská statistika nepředstavuje jednotný směr. Berger (2000, s. 1271-1273) například rozděluje bayesovskou statistiku na několik přístupů: objektivní, subjektivní, robustní, frekvenčně-bayesovský a kvazi-bayesovský.

⁴ Přednáška na toto téma od neurovědyce Daniela Wolperta dostupná na http://www.ted.com/playlists/1/how_does_my_brain_work. V neurovědě smyslové vjemy představují data a apriorní informace představují naše představy o světě, které získáváme na základě zkušenosti.

hodnotě hladiny významnosti jsou i ve frekvenční statistice zcela na výzkumníkovi a jsou tedy také subjektivní⁵.

Vzhledem k přítomnosti subjektivity ve všech vědeckých snahách o analýzu dat, by si měl badatel zvolit takové inferenční paradigma, které nabízí nejotevřenější předložení předpokladů modelu. Takovým přístupem je bayesovský přístup, neboť apriorní rozdělení je jasně specifikovaným a jasně stanoveným předpokladem modelu. Naopak frekvenční modely jsou zřídka prezentovány s upozorněními jako „Varování: vědecké závěry zde prezentované předpokládají opakované pokusy, které se nikdy nestaly.“ nebo „Pozor: apriorní předpoklady tohoto modelu nejsou diskutovány nebo objasněny.“ Základní vědecká zásada při publikování empirických zjištění je, že všechny důležité charakteristiky modelu by měly být čtenáři poskytnuty. Je jasné, který z přístupů je z hlediska tohoto kritéria více „vědecký“ (Gill, 2008, s. 27).

Gill (2008) uvádí několik důvodů, proč se podle něj zvedá zájem o bayesovskou statistiku: Za prvé, společnost zásadně zvýšila poptávku po statistické analýze všeho druhu a dnešní dobu proto lze považovat za zlatý věk statistiky. Za druhé, mnoho vědců, kteří zhodnotili dostupná paradigma, zjistilo, že alternativy bayesovského přístupu jsou ve svém jádru nekonzistentní a nedokonalé. Za třetí, po nedávném průlomu ve statistickém programování bylo možné snadno tvořit bayesovské statistické modely (Gill, 2008, s. 1-2).

2.4 Vhodnost bayesovské statistiky pro sociologická data

Data, se kterými se v sociálních vědách běžně setkáváme, jsou vhodná pro bayesovský způsob takovým způsobem, jaký tradiční statistická analýza nemůže nabídnout (Gill, 2008, s. 3). Jak už bylo zmíněno, data, která v sociologii získáváme, nemají experimentální povahu. Zároveň v sociologii neprovádíme mnoho opakovaných výběrů z populace, ale obvykle jen jeden výběr. Bayesovská statistika je ideálním přístupem pro usuzování v sociálních vědách, neboť v nich je žádoucí vytvářet modely testující teorie s nově naporozovanými daty, které jsou založené na předešlých výzkumech či znalostech. Nikdy nezačínáme analýzu dat s absolutně žádnou předchozí (apriorní) představou o zkoumaném jevu (nebo bychom přinejmenším neměli) (Gill, 2008, s. 18).

⁵ Vzhledem k tomu může být bayesovský subjektivismus „objektivnější“ než frekvenční objektivismus.

2.5 Omezení

Navzdory filozofickým rozdílům mezi frekvenčním a bayesovským způsobem uvažování existují dvě situace, kdy analýzy provedené těmito přístupy povedou ke stejným výsledkům. Jednak půjde o situace, kdy vložená apriorní informace bude velmi neinformativní (podrobnější popis neinformativních apriorních rozdílů v podkapitole 6.1.3.3). Druhý případ, kdy budou výsledky shodné, je situace, když budeme analyzovat data s velkým množstvím pozorování. V takové situaci je apriorní informace převážena daty a výsledky obou přístupů se budou opět shodovat (Gill, 2008, s. 4). Přesto, že v takových případech budou výsledky obou přístupů shodné, je lepší zvolit bayesovský přístup, neboť „(a) *symptotické* vlastnosti výběrových charakteristik sice při velkých výběrech z výpočetního hlediska eliminují zásadní rozdíly mezi oběma přístupy, ale interpretační potíže a metodologické nejasnosti s klasickými odhady a testy přetrvávají...“ (Hebák, 2012a, s. 73; zvýraznění v originále).

3. Historie statistiky

Zde stručně shrneme historii statistiky. V této vědní disciplíně téměř celé 20. století dominovala frekvenční statistika. To je také důvodem, proč i dnes v aplikacích i ve výuce přetrvává. Bayesovský přístup je historicky starší, ale díky jeho přednostem v současné době zažívá velký rozkvět, který byl některými předními statistiky předpovídán (Hebák, 2012a, s. 70).

3.1 Počátky statistiky

Pod označením statistika se skrývají dva vědecké obory – popisná statistika, kterou lze zaznamenat už ve starověku, a matematická statistika, která se zrodila ve 20. století. Popisná (deskriptivní) statistika se zabývá původním předmětem, a to sběrem informací o státu. V tomto smyslu se statistika používala už ve starověku (např. Sumerská říše, Egypt, Řím), kdy se provádělo sčítání lidu za účelem vybírání daní. Záznamy se vedly i o počtu kusů domácího zvířectva a úrodě apod. Ve středověku se pak statistická šetření omezovala pouze na soupisy majetku (Žák, 2006, s. 1-2).

3.2 Politická aritmetika 16. a 17. století

V 16. a 17. století se v Anglii začíná uplatňovat tzv. politická aritmetika (v podstatě dnešní demografie), která se zabývala údaji o narozeních a úmrtích, na jejichž základě se pokoušela určit vývoj obyvatelstva. Za zakladatele této disciplíny je považován John Graunt (1620-1674). Vliv a rozvoj základních teoretických koncepcí

pravděpodobnosti se začal projevovat ve statistice stále výrazněji. Významnými osobnostmi v tomto období byl také Edmund Halley (1656-1742) a Adolphe Lambert Quételet (1796-1874), který vypracoval zásady moderních sčítání lidu a který také jako první objevil normální rozdělení, když na základě velkého souboru dat vypočítal rozměry průměrného člověka. Normální rozdělení podle Quételeta znamenalo, že se příroda snaží vytvořit ideálního člověka, avšak v různé míře chybuje (Reiterová, 1998, s. 75-77). Quételet je považován za duchovního otce pojmů průměr, střední hodnota, rozptyl a rozdělní (Žák, 2006, s. 2-3).

3.3 Subjektivně orientovaná 2. polovina 18. – 19. století

Druhá polovina 18. a téměř celé 19. století bylo z hlediska statistické metodologie spíše subjektivně/bayesovsky orientované. Za zakladatele bayesovského přístupu je považován Thomas Bayes (1701-1761), podle něhož je také pojmenován (Hebák, 2012a, s. 72).

Thomas Bayes byl presbyteriánský duchovní, který studoval logiku a teologii na Univerzitě v Edinburghu. Přestože byl reverendem, zajímal se o matematiku, přírodní vědy a nebeskou mechaniku (Bellhouse, 2004, s. 21). Jeho slavná práce *An Essay towards Solving a Problem in the Doctrine of Chances* (Bayes, 1763) vyšla dva roky po jeho smrti. V této práci Bayes jako první použil pravděpodobnost induktivně a založil matematické základy pro pravděpodobnostní usuzování. Pravděpodobnostní pravidlo, které v tomto článku jako první formuloval, se po něm dnes nazývá Bayesovo pravidlo. Toto pravidlo umožňuje zkombinovat výzkumníkovy teorie s empirickými pozorováními (Gill, 2008, s. 21).

Bayesovo pravidlo později rozšířil Pierre Simon Laplace (1749-1827) a mnoho dalších následovníků (Haigh, 2012, s. 34). Jeho rozvíjení probíhalo kvůli obtížnosti výpočtů především v teoretické rovině. To se s novými výpočetními a simulačními technikami v 90. letech 20. století změnilo, což umožnilo aplikaci Bayesova pravidla na empirická data.

3.4 Objektivisticky orientovaný konec 19. – 20. století

Od konce 19. století a téměř celé 20. století byla doba velkého rozvoje objektivistické/klasické statistické školy. V tomto období tento přístup jasně dominoval. Hlavními představiteli klasické doby byly např. Karl Pearson, Egon Pearson, Jerzy Neyman, William Sealy Gosset (známý jako Student) a především Ronald Alymer Fisher (Hebák, 2012a, s. 72).

Statistika se v tomto období snažila především shromáždit mnoho dat od co největšího počtu respondentů. Cílem tohoto snažení bylo obsáhnout v rámci šetření co největší část populace, aby tak mohla získat co nejpřesnější obraz stavu společnosti. Náročnost takovýchto šetření vedla postupně statistiky k otázce, zda je skutečně potřeba zkoumat celou populaci nebo stačí zkoumat pouze její reprezentativní výběr. Na základě této myšlenky a díky propojení s počtem pravděpodobnosti se na počátku 20. století zrodila nová statistická disciplína – matematická statistika (Žák, 2006, s. 3).

Největší postava objektivního přístupu ve statistice (někdy také nazývané fisherovské statistice) R. A. Fisher (1890-1962) striktně odmítal jakýkoliv jiný než objektivistický přístup – byl tedy velkým odpůrcem subjektivistické školy. Přestože si Bayese údajně vážil pro jeho matematický výklad, bayesovský přístup odmítal (dokonce si byl jist tím, že Bayes svou esej za svého života nevydal záměrně, neboť o ní sám pochyboval) (Edwards, 2004, s. 36-37). O podmíněné pravděpodobnosti (která je základem bayesovského přístupu) psal jako o paradoxu, „... který je jako neproniknutelná džungle a zastavuje rozvoj směřující k vytříbení statistických konceptů.“ (Fisher, 1922, s. 311) Údajně také pracoval na tom, aby inverzní pravděpodobnost a bayesiánství zdiskreditoval. Není bez zajímavosti, že Fisher později vyvinul tzv. fiduciární usuzování (*fiducial inference*), které se ukázalo být chybné a které, jak později dokázal Lindley (1958), je konzistentní, pouze pokud je použito tak jako bayesovský přístup s uniformním apriorním rozdělením (Gill, 2008, s. 21-22). Někteří statistikové (dodejme, že jde většinou o zastánce bayesovského přístupu) dokonce spekulují nad tím, zda nebyl Fisher ve skutečnosti bayesiánem⁶.

Nutno podotknout, že počátek používání statistiky v sociologii se datuje právě do tohoto období. Přestože sociologie jako věda vznikla v 2. polovině 19. století a kvantitativní metody byly v sociologii používány už od jejího vzniku, na počátku byly založené spíše na útržkovitých datech a použité statistické metody byly pouze jednoduché a deskriptivní (Raftery, 2000, s. 654). Statistika se v sociologii začala uplatňovat od poloviny 40. let 20. století. Raftery (2000) rozlišuje tři období podle toho, do jaké míry a jakým způsobem byla statistika v sociologii používána. První generace metod v období od poloviny 40. let do konce 50. let 20. století se zabývala především tvorbou kontingenčních tabulek z výzkumů o malém počtu diskrétních proměnných.

⁶ Například podle de Finettiho názoru Fisher nejspíš cítil potřebu vyjádřit své závěry bayesovsky, ale chtěl k nim dojít z opačné strany (frekvenčně) (de Finetti, 1975, s. 248).

Typickým příkladem tohoto období je výzkum sociální mobility (Raftery, 2000, s. 655). Základním textem tohoto období je například Lazarsfeldova kniha *Mathematical Thinking in the Social Sciences* (1954) (Edling, 2002, s. 198). Druhá generace metod přišla na začátku 60. let 20. století a zkoumala data v podobě mnoha proměnných zjištěných o jedné jednotce (Raftery, 2000, s. 655). Důležitou publikací a dnes již klasickým textem byla v tomto období Colemanova kniha *Introduction to Mathematical Sociology* (1964) (Edling, 2002, s. 198). Třetí generace metod zabývající se složitějšími proměnnými i statistickými metodami přišla na konci 80. let 20. století a trvá dodnes (Raftery, 2000, s. 655).

3.5 První renesance bayesovského přístupu v 2. polovině 20. století

Přes kritiku nejvlivnějšího statistika 20. století se bayesovská statistika začala znovu probouzet, a to přibližně v polovině 20. století. Následovalo asi dvacet let bouřlivých diskusí o oprávněnosti či neoprávněnosti tzv. neobayesiánství (Hebák, 2012a, s. 73). Boj mezi oběma přístupy byl nejvypjatější v druhé polovině 20. století, diskuse však pokračuje dodnes (Rosentah, 2008, s. 247).

Bayesovská škola nicméně postupně sílila, o což se zasloužili např. Bruno de Finetti, Leonard Jimmy Savage, Harold Jeffreys, Jose Miguel Bernardo, Dennis Lindley, a další. Tito vědci znovuobnovili zájem o bayesovský přístup, a to jako odpověď na pozorované nedostatky v klasických technikách. Přes vynikající teoretickou práci těchto statistiků bylo stále nemožné nebo jen velmi obtížné složitější modely spočítat. Naštěstí tento problém je již také překonán díky revoluci ve statistických výpočetních technikách v 90. letech 20. století, která vedla k druhé renesanci bayesovského paradigmatu (Gill, 2008, s. 21-22).

Ve vývoji bayesovského přístupu považuje Hebák (2012a) za zlomový rok 1994, kdy známý statistik Maurice Kendall zařadil do své všeobecně uznávané statistické řady publikací *The Advance Theory of Statistics* díl s názvem *Bayesian Inference* (autorem je Anthony O'Hagan). Dnes všechny významné statistické časopisy publikují články o bayesovské problematice, případně vycházejí celá čísla věnovaná bayesovské statistice (Hebák, 2012a, s. 73).

3.6 Druhá renesance bayesovského přístupu v 90. letech 20. století

3.6.1 MCMC algoritmy a prudký rozvoj

Vzestup zájmu o bayesovský přístup začal v 90. letech 20. století po objevení metody Monte Carlo⁷ zkombinované s Markovovými řetězci (v angličtině *Markov Chain Monte Carlo*, ve zkratce MCMC algoritmy/metody). (Podrobněji se MCMC algoritmy budeme zabývat v části 5.2.3.)

Od té doby získávají bayesovské metody ve statistice stále důležitější postavení (Gill, 2008, s. xix). MCMC algoritmy vyřešily přetrvávající problém v bayesovské analýze – bayesovské modely byly dříve analyticky neřešitelné, neboť vedly k multidimenzionálním integrálním počtům (Gill, 2008, s. 22). Prudký rozvoj MCMC metod na přelomu 20. a 21. století vedl k velkému boomeru v používání bayesovské statistiky, která se tak postupně rozšířila do téměř všech vědních oborů (Hebák, 2012a, s. 74). Zatímco před 20 lety se o bayesovskou statistiku zajímali pouze vybraní odborníci, dnes tato oblast statistiky zažívá obrovský nárůst zájmu. Ve statistických časopisech jsou již běžné teoretické i aplikované články, vychází mnoho knih poskytující praktické návody na bayesovskou analýzu (jmenujme např. Lynch, 2007; Bolstad, 2007; Hoff, 2009; Kruschke, 2011). Bayesovská statistika začíná být i tématem statistického vzdělávání (Western, 1999, s. 7).

Berger (2000) zkoumal vývoj počtu publikací o bayesovském přístupu. Během prvních 200 let od doby, kdy byl tento přístup vymyšlen (mezi lety 1769-1969) bylo o bayesovské statistice napsáno přibližně 15 knih. Během následujících dvaceti let (mezi lety 1970-1989) odhaduje, že vyšlo přibližně 30 knih. Během následujících deseti let (mezi lety 1990-1999) bylo napsáno zhruba 60 bayesovských knih, vedle toho probíhalo v těchto letech mnoho bayesiánských konferencí a vycházelo mnoho článků.

⁷ Proč Monte Carlo? Nemá to nic společného s kasinem ve stejnojmenné části Monackého knížectví, ve skutečnosti šlo původně o kódové označení těchto metod ve vojenských záležitostech. (Haigh, 2012, s. 85) Metoda Monte Carlo byla vyvinuta za druhé světové války v rámci projektu *Manhattan*, jehož cílem bylo vytvořit první atomovou bombu (Rosentah, 2008, s. 208).

Vzniklo také mnoho bayesovských společností z nichž největší jsou *International Society of Bayesian Analysis*⁸ a *The Bayesian Statistical Science Section of the ASA (American Statistical Association)*⁹.

3.6.2 Výuka statistiky – bayesovsky nebo frekvenčně?

V současné době se vedou mezi statistiky diskuse na téma výuky statistiky. Obecně se ukazuje, že zejména ve společenskovědních oborech je statistika vnímána studenty spíše negativně (Davis a Sandifer-Stech, 2006; Collins a Onwuegbuzie, 2007). Někteří zastánci bayesovské statistiky, kteří vyzdvihují bayesovský přístup jako lépe pochopitelný a snadněji interpretovatelný, zastávají názor, že bayesovský přístup by se měl učit jako první, tedy před frekvenčním přístupem (Hebák, 2012a, s. 73). Zatímco v zahraničí se statistici dohadují, kdy studenty s jedním či druhým přístupem seznámit, Hebák (2012a) apeluje na společenskovědní vysoké školy a vyučující statistiky v České republice, aby studenti dostali, alespoň v rámci volitelných kurzů, možnost se s tímto způsobem a jeho přednostmi vůbec seznámit. „Na rozdíl od mnohých vyspělých zahraničních univerzit přetrvává v České republice tradiční výuka statistických úsudků bez jediné změny či aspoň zmínky o existenci bayesovské statistiky.“ (Hebák, 2012a, s. 71) Dodejme, že FSV UK v tomto směru tvoří světlou výjimku.

3.6.3 Situace v ČR

Jak už bylo naznačeno v části o výuce statistiky, bayesovský přístup je v České republice téměř neznámý. O bayesovské statistice toho tedy v češtině zatím mnoho napsáno nebylo. Z hlediska obecných informací o bayesovském přístupu jsou dostupná starší skripta *Bayesovské metody* (Hušková, 1985) a *Texty k bayesovské statistice* (Hebák, 1999)¹⁰.

Zatím posledním českým příspěvkem je Hebákův trojdílný článek nazvaný *Srovnání klasické a bayesovské pravděpodobnosti a statistiky*. V nich se Hebák postupně zabývá porovnáním objektivního a subjektivního chápání pravděpodobnosti, bodovými odhady, intervalovými odhady a testováním hypotéz.

V závěru svého článku Hebák porovnává situaci z hlediska vnímání obou přístupů v České republice a v zahraničí. Zatímco v zahraničí je bayesovský přístup

⁸ <http://bayesian.org/>.

⁹ <http://community.amstat.org/SBSS/Home>.

¹⁰ Na Slovensku je situace podobná, vyšla zde skripta *Bayesovská štatistika* (Pázman, 2003).

všeobecně uznáván a v současnosti existují desítky bayesovsky orientovaných monografií a obrovské množství odborných článků, v České republice nebyla zatím knižně vydána žádná bayesovská monografie, ani článek zaměřující se na odlišnosti klasické a bayesovské školy (trojdílný článek Hebáka je prvním). Ve výuce i aplikacích u nás jednoznačně dominuje klasický přístup, který podle Hebáka patří „...svou výhradností (bez přehánění) spíše do třicátých až sedmdesátých let minulého století.“ (Hebák, 2012c, s. 85)

3.6.4 Budoucí vývoj

Předpovědi o budoucím vývoji by se mezi statistiky zřejmě lišily podle toho, jaký přístup jim je bližší. Například Gill (2008) souhlasí s tvrzením uznávaného odborníka na bayesovskou analýzu Bruna de Finetti, který předpověděl, že do roku 2020 dojde k paradigmatické změně k bayesiánství (Gill, 2008, s. xix) Jsou však také názory, které předpovídají jakési sloučení obou nyní soupeřících přístupů. Hebák jmenuje významného současného bayesovského statistika Bradleyho Efrona, který tvrdí, že pro budoucí vývoj statistiky je podstatné propojení toho lepšího z obou přístupů. „Celkově je možné shrnout, že souboje mezi klasickým a bayesovským přístupem přestávají být nesmiřitelné, přičemž se předpokládá, že všichni statistici aspoň částečně znají přednosti i slabiny obou škol. Hledají se styčné body, které by ve svém důsledku znamenaly posun v oblasti statistické teorie a poskytly lepší metody a postupy pro výzkumníky využívající statistické nástroje při analýze dat.“ (Hebák, 2012c, s. 86)

4. Základní pojmy a koncepty

Cílem statistické inference neboli statistického usuzování je odvodit z dat získaných z výběrového vzorku závěry o celé populaci, pro kterou data nemáme (Iversen, 1984, s. 7). Jak již bylo zmíněno v kapitole o historii statistiky, usuzování na základě dat o výběru se provádí proto, že sběr takových dat je organizačně, časově i finančně méně náročný. Jisté je, že informace z výběru nebudou nikdy stejné, jako bychom získali na základě dat z celé populace (Iversen, 1984, s. 8).

Obecně v rámci statistického usuzování provádíme buď metody odhadování, nebo postupy založené na statistických testech. U odhadování jde o určení hodnot neznámého parametru, k tomu slouží bodový nebo intervalový odhad. Testování hypotéz nám poskytuje podporu pro přijetí nebo odmítnutí hypotézy o parametru nebo pravděpodobnostním rozdělení (Hendl, 2012, s. 172).

4.1 Pravděpodobnost

Statistická inference je založená na teorii pravděpodobnosti, takže je dobré nejprve shrnout základní pravidla pro počítání pravděpodobnosti, abychom pochopili konceptuální základy obou porovnávaných přístupů.

4.1.1 Základní vlastnosti

Vzhledem k tomu, že v následujícím textu se budeme o pravděpodobnost opírat, je důležité nejdříve uvést její základní pravidla. Pravděpodobnost nějakého abstraktního jevu A označujeme jako $p(A)$. Základními vlastnostmi pravděpodobnosti jsou:

1. Pravděpodobnost jevu, který je jistý, se rovná 1.
2. Pravděpodobnost jevu nemožného je rovna 0.
3. Lze-li náhodný jev rozložit na několik vzájemně se vylučujících jevů, pak se jeho pravděpodobnost rovná součtu pravděpodobností těchto jevů (Hendl, 2012, s. 124).

Tímto formálním matematickým určením pravděpodobnosti se řídí všechny typy pravděpodobností. Vychází z nich tedy jak klasická/frekvenční, tak bayesovská škola (Hebák, 2012a, s. 75).

4.2 Typy pravděpodobností a jejich využití

Jak již bylo uvedeno, v jádru rozdílů mezi různými přístupy statistické inference jsou různé definice pravděpodobnosti. V této části se proto budeme různými definicemi pravděpodobnosti zabývat, přičemž se zaměříme především na objektivistické/frekvenční a subjektivistické/bayesovské pojetí.

4.2.1 Klasická (Laplaceova) definice pravděpodobnosti

Klasická definice pravděpodobnosti slouží k výpočtu pravděpodobnosti jevů v situacích, kdy známe všechny možné výsledky náhodného pokusu a víme, že tyto jevy (výsledky) jsou stejně možné. Pravděpodobnost jevu A pak spočítáme jednoduše jako poměr počtu výsledků příznivých jevu A k počtu všech možných výsledků (Hebák a Kahounová, 1994, s. 13). Pravděpodobnost nastoupení nějakého jevu můžeme spočítat před provedením pokusu na základě výše zmíněné logické úvahy.

4.2.2 Statistická/frekvenční definice pravděpodobnosti

U statistické definice pravděpodobnosti vycházíme ze situace, ve které během pokusu lze dostat různé výsledky, a přitom nelze předem určit, který z výsledků získáme. Tento pokus přitom lze libovolně mnohokrát opakovat, aniž by se jednotlivá opakování vzájemně ovlivňovala (Hendl, 2012, s. 123). U statistické definice

pravděpodobnosti můžeme pravděpodobnost nastoupení jevu spočítat až na základě výsledků již provedených pokusů (Hebák a Kahounová, 1994, s. 14). Na základě frekvenční pravděpodobnosti Neyman s Pearsonem vytvořili systém klasické statistické inference s testováním hypotéz a odhady pomocí intervalů spolehlivosti (Iversen, 1984, s. 9). Tento přístup je velmi užitečný. Jeho nevýhodou však je, že často není možné získat velký počet výsledků z přesně stejného jevy-generujícího systému (*event-generating system*) (Gill, 2008, s. 6-7).

Statistická definice pravděpodobnosti se někdy označuje jako frekvenční (četnostní) pravděpodobnost nebo objektivní pravděpodobnost. Frekvenční proto, že předpokládá velký počet opakovatelných nezávislých náhodných pokusů. Při velkém počtu pokusů „... lze *relativní četnost jevu A* považovat za odhad pravděpodobnosti jevu *A*. Vychází se z toho, že dochází-li s růstem počtu pokusů ke *stabilizaci* relativní četnosti jevu *A* kolem nějakého čísla, pak lze toto číslo považovat za odhad pravděpodobnosti jevu *A*.“ (Hebák, 2012a, s. 75; zvýraznění v originále) Tato definice se opírá o dílčí větu zákona velkých čísel, která říká, že „... při dlouhodobě nezměněných podmínkách nezávislých pokusů konverguje relativní četnost jevu *A* podle pravděpodobnosti k pravděpodobnosti tohoto jevu.“ (Hebák, 2012a, s. 75; zvýraznění v originále) Jinými slovy dle teorie se při velkém počtu pokusů bude relativní četnost přibližovat k pravděpodobnosti výsledků spočítané podle klasické definice (Haigh, 2012, s. 4). Platí, že čím více dat posbíráme (např. čím vícekrát hodíme kostkou), tím lepší bude náš odhad. Označení objektivní vychází z toho, že pravděpodobnost nějakého jevu vnímá „... jako *objektivně danou vlastnost tohoto jevu*, bez ohledu na to, zda ji umíme či neumíme určit.“ (Hebák, 2012a, s. 76; zvýraznění v originále)

4.2.3 Subjektivní/bayesovská definice pravděpodobnosti

Bruno de Finetti (1906-1985), jeden z nevlivnějších moderních statistiků (a velký zastánce bayesovské statistiky) napsal tento často citovaný výrok: „PRAVDĚPODOBNOST NEEEXISTUJE.“ (de Finetti, 1974, s. x) Tímto výrokem chtěl říci, že odmítá absolutní pravděpodobnostní tvrzení. Podle něj je každý pravděpodobnostní výrok pouhým vyjádřením názoru, založeném například na zkušenosti nebo znalosti, který je možné změnit na základě nových informací (Haigh, 2012, s. 6).

Subjektivní pojetí pravděpodobnosti se zcela liší od klasického a frekvenčního pojetí. Na rozdíl od frekvenčního pojetí pravděpodobnosti, které vnímá pravděpodobnost jako objektivní danost, bayesovské pojetí pravděpodobnosti vnímá pravděpodobnost jako osobní postoj k této vlastnosti (Hebák, 2012a, s. 77). V subjektivním přístupu používáme termíny jako stupeň důvěry či přesvědčení, osobní pravděpodobnost či subjektivní pravděpodobnost. K jejímu vyjádření používáme také hodnoty od 0 do 1 (neboli od 0 % do 100 %), kde 0 znamená, že není možné, aby daný jev nastal, naopak 1 znamená, že jev určitě nastane (Haigh, 2012, s. 7).

„Bayesovská definice pravděpodobnosti je skutečně **subjektivní** a představuje **stupeň přesvědčení** jednotlivce (běžně označovaný za *degree of belief*), že daný jev nebo daná situace nastane.“ (Hebák, 2012a, s. 77; zvýraznění v originále) Vzhledem k tomu, že pravděpodobnost je v tomto případě skutečně subjektivní, je zřejmé, že různí výzkumníci mohou mít na pravděpodobnost nějakého jevu různé názory. Jak už bylo zmíněno výše, i na subjektivní pravděpodobnost se vztahují základní pravidla pravděpodobnosti, subjektivní pravděpodobnost tedy musí být také vnitřně konzistentní (koherentní) (Haigh, 2012, s. 12).

Lidé přirozeně o věcech přemýšlejí pravděpodobnostně, stejně tak existuje určitá „přirozená konzistence“. Například, když si člověk myslí, že jev A nastane pravděpodobněji než jev B a že jev B nastane pravděpodobněji než C, pak si tento člověk logicky nemůže myslet, že C nastane pravděpodobněji než A (tranzitivita). Subjektivní pravděpodobnost je úzce spjata s teorií rozhodování (*decision-making*) a principem maximalizace osobního užitku (Gill, 2008, s. 7).

4.2.3.1 Je subjektivita vědecká?

Mohlo by se na první pohled zdát, že subjektivní definice pravděpodobnosti je málo vědecká. Ve skutečnosti však vědci velmi často vedou diskuse o odlišných teoriích, aniž by věděli, která z nich je správná. Mají také jenom vysoký stupeň přesvědčení v určitou teorii (Hebák, 2012a, s. 77-78).

Subjektivita hraje ve skutečnosti velký význam ve vědeckých objevech. Vědecká práce neformálně postupuje tak, že vychází z dostupných/předchozích informací o daném fenoménu, které postupně doplňuje o nové objevy. V knize *The Subjectivity of Scientist and Bayesian Approach* (Press a Tanur, 2001) autoři ukazují, že věda je ve skutečnosti kombinací subjektivního a objektivního pohledu. Popisují

příklady dvanácti nejznámějších vědců historie¹¹, jejichž velké vědecké objevy byly na počátku neformálně ovlivněny jejich subjektivní vírou či intuicí.

4.2.4 Využití jednotlivých přístupů

Každý přístup k pravděpodobnosti má své využití. Klasický přístup je omezen na okolnosti, které mají konečný počet výsledků, přičemž každý z nich je stejně pravděpodobný.

Frekvenční přístup je vhodný pro nezávislé jevy, které je možné teoreticky nekonečně opakovat za stejných podmínek. Zde ovšem nastává problém, že ať se snažíme sebevíc, experimentální podmínky nemohou být nikdy absolutně identické a jakákoli limitní hodnota může být pouze odhadována. To znamená, že se týká jen hromadných náhodných jevů. Možnosti jeho využití jsou tedy poměrně úzká (Hebák, 2012a, s. 76).

U mnoha jevů okolnosti nastoupení jevu nastávají pouze jednou a jeho alternativy nemohou být redukovány na nějaký konečný počet stejně pravděpodobných výsledků. V takových případech nám klasický ani frekvenční přístup nemohou sloužit a nejvhodnější variantou pro určení pravděpodobnosti se stává subjektivní přístup. Nicméně, jak Haigh (2012) poznamenává, je vždy dobré mít subjektivní závěry podpořené argumenty klasického nebo frekvenčního přístupu (Haigh, 2012, s. 11-12). Subjektivní pravděpodobnost má tu vlastnost, že pomocí ní můžeme vyjádřit také pravděpodobnosti událostí, které nemohou být opakovány. V tomto smyslu jsou subjektivní pravděpodobnosti obecnější než frekvenční pravděpodobnosti, neboť mohou být použity k vyjádření nejistoty, kterou máme o jedné jedinečné události. Zároveň subjektivní pravděpodobnosti nejsou unikátní, neboť dva lidé mohou mít různé dostupné informace, a tudíž se budou i jejich subjektivní pravděpodobnosti o nějaké události či jevu lišit (Iversen, 1984, s. 9).

Vezměme si například otázku: „Jaká je pravděpodobnost, že v České republice během následujícího roku spadne letadlo?“ nebo „Jaká je pravděpodobnost, že na Ukrajině vypukne válka?“ Z hlediska frekvenčního přístupu jsou takové otázky zcela nepřijatelné, neboť v jedinečných případech nelze stanovit relativní četnost. Tím se dostáváme k bayesovské pravděpodobnosti, kterou v takových situacích použít můžeme

¹¹ Aristoteles, Galileo Galilei, William Harvey, Sir Isaac Newton, Antoine Lavoisier, Alexander von Humboldt, Michael Faraday, Charles Darwin, Louis Pasteur, Sigmund Freud, Marie Curie, Albert Einstein.

(Hebák, 2012a, s. 76-77). Frekvenční pravděpodobnosti nám nemohou dát odpovědi na události v minulosti, neopakovatelné události ani události v budoucnosti. Z toho plyne, že nemohou odpovědět na mnoho otázek a předmětů zájmu sociálních vědců (Jackman, 2009, s. xxxii).

Praktická část

5. Frekvenční vs. bayesovská statistika

V této části se nejdříve zaměříme na postup klasické a bayesovské inference. Na začátku krátce připomeneme postup klasické inference a poté podrobněji rozebereme postup bayesovské inference.

5.1 Frekvenční statistika

Frekvenční statistika je nejrozšířenějším způsobem statistického usuzování, jehož dominance plyne zejména z jejího významného rozvoje v průběhu 20. století (viz. část 3. o historii statistiky). Téměř veškerá statistická inference používaná dnes v sociálních vědách je založená na tomto pojetí statistiky (Iversen, 1984, s. 9). Jak už bylo uvedeno, frekvenční statistika vychází ze statistické či frekvenční definice pravděpodobnosti. Předpokládá tedy velký počet opakovatelných nezávislých náhodných pokusů, jejichž relativní četnost je při velkém počtu pokusů považována za pravděpodobnost daného jevu. Vzhledem k této definici se u frekvenční statistiky předpokládá, že data jsou získávána náhodným výběrem nebo metodou znáhodněného experimentu. Jen na základě takovýchto dat je možné provádět zobecňující úsudky v celé populaci (Hendl, 2012, s. 121).

Otázkou je, zda jsou data v sociálněvědních výzkumech (a v jiných vědních oblastech, kde je frekvenční statistika používána) sbírána na základě náhodného výběru. Sociologická šetření obvykle na pravděpodobnostním výběru trvají, v aplikovaném výzkumu se používají především kvótní výběry. Klasická statistika byla původně vyvinuta pro experimentální uspořádání, dnes je však často používána na data z kvaziexperimentů či pseudoexperimentů, kam spadají i sociologické průzkumy (Soukup, 2010, s. 79). V případech, kdy provádíme frekvenční statistické úsudky na základě jiné metody výběru, validita statistického usuzování a závěry z ní plynoucí lze snadno zpochybnit (Hendl, 2012, s. 172-173).

Frekvenční statistika při výpočtech vychází z tzv. výběrového rozdělení, což je pro ne-statistiky poměrně složitý pojem. Obecně můžeme rozlišit mezi rozdělením výběru (*sample distribution*), což je rozdělení hodnot v jednom konkrétním (již pořízeném) výběru, a výběrovým rozdělením (*sampling distribution*), což je pojem, který se „... neopírá o pravděpodobnosti **vycházející z daného konkrétního výběru**, ale o hůře představitelnou situaci **všech možných výběrů z dané populace**.“ (Hebák,

2012a, s. 73; zvýraznění v originále) Statistika spočítaná frekvenčně na základě výběrového rozdělení ve skutečnosti nemá nic společného s daty získanými z pořízeného náhodného výběru, ale je stanovena předem. Četnostní škola, na rozdíl od bayesovské, tento přístup považuje za rozumnou aproximaci reality (Hebák, 2012b, s. 79).

Zmíněné předpoklady jsou kamenem úrazu ve chvíli, kdy neprovádíme mnoho výběrů z populace, ale pouze jeden výběr, jehož charakteristiky nás zajímají. To způsobuje problémy zejména v interpretaci bodových a intervalových odhadů. Ty podrobněji je rozebereme v další části (v podkapitolách 6.2 a 6.3). U testování hypotéz je dnes standardním měřítkem hladina významnosti 5 %¹². Hodnota 5 % je přitom zcela libovolná. Tato hladina významnosti byla zvolena R. A. Fisherem jako „... číslo patřičně malé i matematicky vhodné...“, sám však připouštěl, že ji můžeme v případě zájmu změnit (Rosentah, 2008, s. 116). Testování hypotéz je jednou z nejvíce kritizovaných oblastí frekvenční statistiky. Podrobněji ji opět rozebereme později (v podkapitole 6.4).

5.2 Bayesovská statistika

Bayesovská inference nám umožňuje za použití nějaké naší původní představy o hodnotě parametru (apriorní informace), a na základě informací z provedeného výběru (tedy našich dat), určit hodnotu tohoto parametru (pomocí posteriorního rozdělení¹³) (Iversen, 1984, s. 11).

Bayesovský přístup při odhadech, usuzování a vytváření modelů postupuje těmito kroky:

1. Na základě výchozí informace specifikujeme apriorní rozdělení pro parametry modelu, které tuto informaci zahrnují.
2. Na základě dat vytvoříme věrohodnostní funkci.
3. Pomocí Bayesova vzorce z apriorního rozdělení a věrohodnostní funkce získáme posteriorní rozdělení pro parametry modelu.
4. Pomocí simulačních technik získáme výběr z posteriorního rozdělení parametrů.

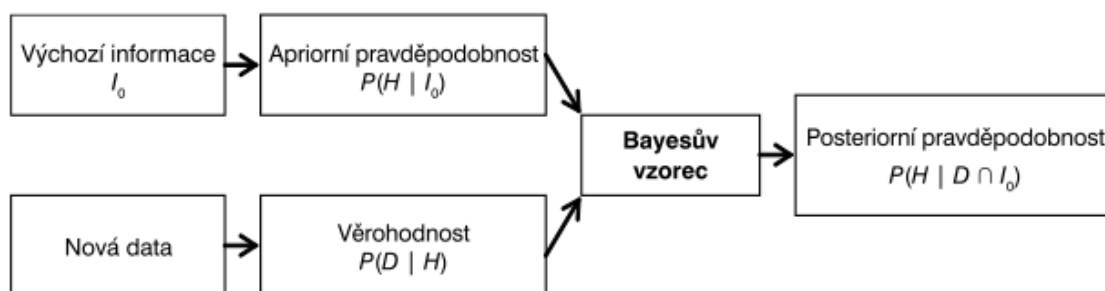
¹² Podle Rosentaha (2008) je pojem statistická významnost ve vědě srovnatelný s právnickým vyjádřením „nade vší pochybnost“.

¹³ Apriorní znamená „nezávislý na smyslové zkušenosti“, zatímco aposteriorní znamená „založený na smyslové zkušenosti“.

5. Na jeho základě vytvoříme pomocí deskriptivní statistiky závěry o tomto modelu.

Tento proces může být s nově získanými daty bez problémů opakován (Lynch, 2007, s. 2; Gill, 2008, s. 5-6). První tři kroky bayesovského přístupu můžeme znázornit v přehledném obrázku (Obrázek 1). Na základě relevantních informací vytvoříme apriorní pravděpodobnost a z dat vytvoříme věrohodnostní funkci. Pomocí Bayesova vzorce z nich spočítáme posteriorní pravděpodobnost.

Obrázek 1. Proces revize pravděpodobností po získání nových dat.



Zdroj: Hebák (2012a, s. 84).

5.2.1 Interpretativní a analytické základy

Obecně lze říci, že bayesovská statistika leží na dvou základech: interpretativním a analytickém. V bayesovském přístupu jsou statistické parametry interpretovány pravděpodobnostně, jinými slovy hodnoty parametrů jsou popsány pravděpodobnostními rozděleními. Analyticky jsou tato pravděpodobnostní rozdělení parametrů nalezena pomocí pravidla podmíněné pravděpodobnosti známého jako Bayesova věta¹⁴ (Western, 1999, s. 8).

5.2.1.1 Interpretativní základy

Zatímco ve frekvenční statistice jsou parametry pravděpodobnostních rozdělení **konstanty** (tj. mají jednu konkrétní, ale neznámou hodnotu), v bayesovské statistice jsou parametry chápány jako **náhodné veličiny**. Pravděpodobnostní rozdělení těchto náhodných veličin je velmi užitečným nástrojem (Hebák, 2012a, s. 77). Toto pravděpodobnostní rozdělení kvantifikuje výzkumníkovu nejistotu o hodnotě parametru. V praxi sociologové často interpretují p -hodnoty a intervaly spolehlivosti jako popisy pravděpodobné hodnoty parametrů. Z toho plyne, že paradoxně většina

¹⁴ Pro Bayesovu větu najdeme v literatuře také označení vzorec, teorém či pravidlo.

sociologů interpretuje frekvenční závěry bayesovsky, ačkoli neprovádějí bayesovské výpočty. Přidělením pravděpodobností hodnotám parametrů je tedy bayesovské usuzování silně sociologicky intuitivní (Western, 1999, s. 9).

Ještě před napozorováním dat a započítím jakékoli analýzy je v bayesovské statistice parametrům přiděleno apriorní rozdělení. Přidělení spočívá ve výběru vhodného typu rozdělení (podle toho, zda je proměnná spojitá či diskrétní) a nastavení jeho parametrů, které ho charakterizují. Apriorní rozdělení je potom použito v kombinaci s daty k odhadu parametru. Pokud známe o zkoumaném jevu jen málo, pak by apriorní rozdělení mělo být vágní, pokud o něm naopak víme mnoho, pak by mělo být vyjádřeno úzce a specificky (Gill, 2008, s. 19). Apriorním rozdělením se budeme konkrétněji zabývat později (v podkapitole 6.1).

5.2.1.2 Analytické základy

Analytické základy bayesovské statistiky jsou zakořeněné v základních pravidlech pravděpodobnosti, které jsou popsány v části 4.1 věnované pravděpodobnosti.

5.2.1.2.1 Podmíněná pravděpodobnost

Centrální pro bayesovské usuzování je představa podmíněnosti (Gill, 2008, s. 8). Podmíněná pravděpodobnost znamená, že pravděpodobnost výskytu určitého jevu závisí na tom, zda nastal, nebo nenastal nějaký jiný jev. Podmíněné pravděpodobnosti se značí $P(A|B)$ (což čteme pravděpodobnost jevu A za předpokladu, že nastal jev B). Pro podmíněné pravděpodobnosti platí všechna základní pravidla pravděpodobnosti. Při úvahách o podmíněné pravděpodobnosti se často používá Bayesova věta (Hendl, 2012, s. 126).

5.2.1.2.2 Bayesova věta

Na základě základních pravidel pravděpodobnosti lze jednoduše odvodit¹⁵ Bayesovu větu, která je základem bayesovské inference:

$$P(A|B) = \frac{P(A) * P(B|A)}{P(B)}$$

¹⁵ Odvození Bayesovy věty najdeme například ve Western (1999, s. 12-13).

Bayesovo pravidlo je pojmenované po jeho autorovi Thomasi Bayesovi, který jeho odvození prokázal¹⁶. Samotná Bayesova věta (někdy se jí říká také zákon o inverzní pravděpodobnosti) je všeobecně uznávaná. Kontroverzní je na ní její aplikace se subjektivním konceptem pravděpodobnosti a nevýběrovou informací v apriorním rozdělení (Western, 1999, s. 12-13). Pomocí Bayesovy věty můžeme na základě relevantních informací aktualizovat naše původní informace vyjádřené jako pravděpodobnosti. Demonstrujme to na následujících příkladech.

5.2.1.2.3 Příklad 1 „Alarm“

Představme si situaci, že nám na mobilní telefon přijde SMS, že se v našem domě spustil bezpečnostní alarm. Než začneme panikařit, vzpomeneme si na Bayesův vzorec a spočítáme si, jaká je pravděpodobnost, že je náš dům vykrádán. Převědeme naši otázku do podmíněné pravděpodobnosti: jaká je pravděpodobnost, že je náš dům vykrádán za předpokladu, že se spustil alarm? Víme, že dle statistik je v Praze pravděpodobnost vykradení rodinného domu 0,9 %¹⁷. Dále víme, že máme v domě nainstalovaný kvalitní bezpečnostní systém, u nějž je výrobcem deklarováno, že pravděpodobnost spuštění alarmu za předpokladu, že dojde k vloupání, je 90 %. Naopak pravděpodobnost, že se chybně spustí alarm, za předpokladu, že k žádnému vloupání nedojde, je pouhé 1 %.

Pro přehlednost:

- $p(\text{vykradení}|\text{alarm}) = ?$
- $p(\text{alarm}|\text{vykradení}) = 0,90$
- $p(\text{alarm}|\text{nevykradení}) = 0,01$
- $p(\text{vykradení}) = 0,009$
- $p(\text{nevykradení}) = 0,991$

Bayesův vzorec pro náš příklad zní:

¹⁶ Bayes byl první, kdo toto pravidlo explicitně vytvořil, ale byl to Laplace, kdo jej (zřejmě nezávisle) podrobněji rozpracoval, a jehož práce je vzhledem k dnešní praxi bayesovské statistiky možná relevantnější (Gill, 2008, s. 10).

¹⁷ Na základě údajů z: http://www.csobpoj.cz/cs/o-spolecnosti/pro-media/tiskove-zpravy/Documents/2012_05_11%20TZ%20Pravd%C4%9Bpodobnost%20vloup%C3%A1n%C3%AD%20podle%20kraj%C5%AF.pdf.

$$p(\text{vykradení}|\text{alarm}) = \frac{p(\text{alarm}|\text{vykradení}) * p(\text{vykradení})}{p(\text{alarm}|\text{vykradení}) * p(\text{vykradení}) + p(\text{alarm}|\text{nevykradení}) * p(\text{nevykradení})}$$

Po dosazení hodnot dostaneme:

$$p(\text{vykradení}|\text{alarm}) = \frac{0,9 * 0,009}{0,9 * 0,009 + 0,01 * 0,991} = \frac{0,0081}{0,0081 + 0,00991} = \frac{0,0081}{0,01801} = 0,44975 \cong 45 \%$$

Výsledná pravděpodobnost nás nejspíš trochu uklidní. Pravděpodobnost, že je náš dům právě vykrádán, za předpokladu, že se spustil alarm (který nás upozornil pomocí SMS) je jen 45 % (oproti 90% spolehlivosti alarmu).

5.2.1.2.4 Příklad 2 „Sedmilháři“

Předpokládejme, že v České republice je část populace už natolik otrávená odpovídáním na výzkumy veřejného mínění, že založila skupinu „Sedmilháři“. Členové této skupiny tvoří 10 % populace a jejich cílem je odpovídat na otázky lživě, aby zkreslili výsledky. Výzkumníci se snaží členy této skupiny odhalit a díky proškoleným tazatelům se daří identifikovat členy této skupiny s přesností 95 %. Respondenti, kteří nejsou členy skupiny (tedy upřímní respondenti) jsou jako nečleni identifikováni s přesností 100 %. Zajímá nás, jaká je pravděpodobnost, že jedinec je skutečně členem skupiny, pokud je pozitivně klasifikován.

- $p(\text{Sedmilháři}|\text{Identifikace}) = ?$
- $p(\text{Identifikace}|\text{Sedmilháři}) = 0,95$
- $p(\text{Identifikace}|\text{Ne-Sedmilháři}) = 1$
- $p(\text{Sedmilháři}) = 0,1$
- $p(\text{Ne-Sedmilháři}) = 0,90$

Pravděpodobnosti dosadíme do Bayesova vzorce a dostáváme:

$$p(\text{Sedmilháři}|\text{Identifikace}) = \frac{0,95 * 0,1}{0,95 * 0,1 + 1 * 0,9} = \frac{0,095}{0,095 + 0,9} = \frac{0,095}{0,995} = 0,095 \cong \cong 10 \%$$

Přestože jsou členové skupiny „Sedmilháři“ identifikováni s přesností 95 %, pomocí Bayesova vzorce jsme zjistili, že když je respondent identifikován jako člen skupiny, ve skutečnosti je členem v 10 % případů.

5.2.1.2.5 Zobecnění Bayesova vzorce na data

Jiné odvození Bayesova vzorce je založené na tom, že sdružená pravděpodobnost dvou jevů, označme je nyní jako P a D, může být zapsána jako součin pravděpodobnosti jednoho z jevů a podmíněné pravděpodobnosti druhého jevu, za předpokladu prvního jevu.

$$P(P \cap D) = P(P) * P(D|P)$$

Při prohození dvou jevů může být pravděpodobnost dvou jevů zapsána také jako:

$$P(D \cap P) = P(D) * P(P|D)$$

Protože levé strany jsou si rovny, pak obě pravé strany jsou si také rovny. Když je dáme do rovnosti a upravíme je, pak dostaneme:

$$P(P|D) = \frac{P(P) * P(D|P)}{P(D)}$$

Zjednodušená Bayesova věta se uvádí s vynecháním $P(D)$, což je prediktivní pravděpodobnost vyplývající z věrohodností a apriorních pravděpodobností (Hebák, 2012a, s. 84). Vynechává se proto, že může být někdy těžce spočitatelná a nemění relativní pravděpodobnost parametru. Zjednodušená Bayesova věta tedy zní:

$$P(P|D) \propto P(P) * P(D|P) \quad , \text{ kde } \propto \text{ znamená „je úměrná k“}.$$

Nyní si můžeme představit, že písmeno P představuje parametr (případně hypotézu nebo model) a písmeno D představuje data (Iversen, 1984, s. 12-13). Levá strana Bayesovy věty $P(P|D)$ představuje posteriorní pravděpodobnost neboli pravděpodobnost parametru P, za předpokladu napozorovaných dat D. Na pravé straně je pravděpodobnost $P(P)$, tedy apriorní pravděpodobnost přidělená parametru před tím,

než nám jsou data známa. Pravděpodobnost $P(D|P)$ je výběrová hustota dat (*sampling density of the data*), které je proporcionální věrohodnosti neboli informaci poskytnuté daty (Western, 1999, s. 13). Slovně tedy můžeme Bayesovu větu vyjádřit jako:

$$\text{Posteriorní rozdělení} \propto \text{apriorní rozdělení} * \text{věrohodnost}$$

Když je jednou posteriorní rozdělení spočítáno prostřednictvím daného vzorce, je o něm všechno známo a je pouze na výzkumníkovi, aby zdůraznil body zájmu. Bayesiáni popisují posteriorní rozdělení pomocí průměru, variance, modu, kvantilů atd., což je praktické, protože tak to dělají nebayesiáni (Gill, 2008, s. 19).

5.2.2 Kritika a reakce na ni

Základní neshody mezi bayesiány a frekvenčními statistiky jsou založené na dvou hlavních bodech. 1) Zda data a hypotézy mohou mít stejný status jako náhodné proměnné. 2) Zda použití apriorní pravděpodobnosti vkládá do analýzy příliš subjektivity (Lynch, 2007, s. 71).

ad 1) Použití pravděpodobnostního rozdělení pro parametry bayesiáni odůvodňují tím, že vycházejí ze subjektivního pojetí pravděpodobnosti. To jim umožňuje vyjádřit vlastní nejistotu o hodnotě parametru. Tento přístup má naopak několik podstatných výhod. Jednak umožňuje jednodušší interpretaci v porovnání s frekvenčním přístupem. Umožňuje také přirozeně vložit do modelu předchozí znalosti o parametru. Bývá zvykem v člancích či statích na začátku shrnout dosavadní znalosti o studovaném problému. Frekvenční přístup oproti bayesovskému nemá žádný nástroj, kterým vložit předchozí znalosti do modelu. Výsledek bayesovské analýzy nám přináší také mnohem více informací o parametru (celé rozdělení, z kterého můžeme různé statistiky spočítat), zatímco u frekvenční statistiky máme k dispozici standardně jen maximálně věrohodný odhad a standardní chybu (Lynch, 2007, s. 71).

ad 2) Subjektivita apriorních pravděpodobností je nejkritizovanějším bodem v bayesovském přístupu, neboť subjektivita je podle mnohých v rozporu s vnímáním vědy jako objektivní a jednoznačné (Rosentah, 2008, s. 5). Na uvážení výzkumníka není jen samotná pravděpodobnost, ale také volba pravděpodobnostního rozdělení a jeho parametrů. Je však třeba mít na mysli, že je to pouze první a výchozí krok. Výhodou je možná změna názoru po získání nových informací, což je možnost pro bayesovskou statistiku typická (Hebák, 2012a, s. 78).

Bayesiáni kontrují tím, že i ve frekvenční statistice je mnoho subjektivních prvků – například hladina významnosti, podle které jsou pak určovány výsledky za statisticky signifikantní nebo nikoliv. Při hodnocení signifikance se také často výzkumníkům doporučuje uvážit, zda je daný statisticky významný výsledek také věcně významný, což je jistě také subjektivní záležitost. Navíc subjektivní část bayesovské analýzy, tedy ta apriorní pravděpodobnost, je transparentní – pokud jsou výsledky založené na nějaké nevěrohodné informaci, můžeme je jednoduše zpochybnit (Western, 1999, s. 11).

Další odpovědí na kritiku subjektivity bayesovské statistiky je, že apriorní rozdělení je zvláště v sociálněvědních výzkumech převáženo daty. Apriorní rozdělení mohou být také úmyslně volena jako neinformativní tak, aby vedla ke stejným výsledkům jako maximálně věrohodné odhady používané ve frekvenční statistice. Vliv apriorní informace může být také vyhodnocen po provedené analýze, kdy zvážíme, zda jsou posteriorní výsledky rozumné. Ohodnocení rozumnosti výsledků, jak už bylo zmíněno, je společným rysem bayesovské i frekvenční analýzy (Lynch, 2007, s. 72).

Podle Hebáka je v současné době kritika bayesovské statistiky „...spíše už jen *umíněností* některých vyučujících i výzkumníků, vyplývající z částečné nebo úplné *neznalosti argumentů* bayesiánů, a někdy dokonce jen *jakoby nadřazenou neochotou* některých *klasiků* respektovat vývoj, ke kterému nesporně došlo.“ (Hebák, 2012a, s. 71-72; zvýraznění v originále) Obecně můžeme říci, že kritika bayesovské statistiky je v zásadě filozofická. Vzestup bayesovské statistiky v posledních desetiletích má naopak pragmatické důvody.

5.2.3 MCMC algoritmy

Jak už bylo zmíněno v kapitole o historii statistiky, složitější bayesovské modely se neobejdou bez MCMC algoritmů neboli Monte Carlo simulací uplatňujících Markovské řetězce. Jejich objevení spustilo velký zájem o bayesovské statistické úsudky v mnoha vědních oborech a v téměř všech oblastech statistiky (Hebák, 2012c, s. 85). V této práci nastíníme jen základní smysl této techniky (více o nich např. Jackman, 2009, Part II; Gill, 2008, kap. 9 a kap. 11-13) Cílem MCMC algoritmů je nahradit složité analytické integrování pomocí iterační práce na počítači (Gill, 2008, s. 30). Umožňují nám z apriorního rozdělení a podmíněné hustoty pravděpodobnosti našich dat namodelovat posteriorní rozdělení námi odhadovaných parametrů.

Metody Monte Carlo jsou simulace prováděné především za pomoci počítače. Simulace provádíme pro prozkoumání chování složitějších modelů a představují často jednodušší cestu, jak odhadnout neznámé pravděpodobnosti. Pokud známe pravděpodobnosti základních jevů a mechanismus, jak nové jevy vznikají z jevů základních, pak můžeme využít proces simulace – proces mnohokrát opakujeme a jednotlivé pravděpodobnosti odhadujeme pomocí relativních četností jevů, které nás zajímají (Hendl, 2012, s. 130).

MCMC metoda je iterativní proces, který generuje série hodnot nazývané řetězce pro každý parametr ve specifikovaném modelu. Teorie říká, že pokud MCMC proces běží dostatečně dlouho (tzn. nagenereje stovky až tisíce hodnot), budou generované hodnoty konvergovat s hodnotami, které bychom získali ze sdruženého rozdělení parametrů. Markovův řetězec představuje algoritmus vytvářející iterativní proces, ve kterém každá generace simulovaných hodnot je bezprostředně závislá na předcházející generaci. Z toho vyplývá, že je nutné nejprve vložit nějaké vstupní hodnoty, které bychom měli volit velmi pečlivě. Jakmile skončí fáze zahoření (*burn-in*), jejíž hodnoty se vyřazují, a je dosaženo konvergence, mohou být simulované hodnoty použity k vytvoření posteriorního rozdělení (Woodward, 2012, s. 8-9). Nejpoužívanějšími MCMC algoritmy jsou Gibbsovo vzorkování (*Gibbs Sampling*) a Metropolisův algoritmus (Gill, 2008 s. 30-37).

6. Porovnávací analýza

Odhady a testování hypotéz v bayesovské statistice nejsou prováděny a formulovány stejným způsobem jako ve frekvenční statistice. Výsledkem bayesovské analýzy je posteriorní rozdělení, jehož popis může zahrnovat hodnoty kvantilů (alternativa bodového odhadu) a kredibilní intervaly (alternativa intervalů spolehlivosti), které lze využít také k testování hypotéz¹⁸ (Gill, 2008, s. 4-5). Obecně lze říci, že „(v)elkou výhodou bayesovského přístupu je skutečnost, že k bodovému i intervalovému odhadu, jakož i pro testování statistických hypotéz, je znalost

¹⁸ I ve frekvenční statistice mezi intervalem spolehlivosti a testem významnosti parametru existuje formální vztah ekvivalence. Interval spolehlivosti lze použít při testování významnosti – „(p)okud hodnota hypotetického parametru neleží v intervalu spolehlivosti, hypotézu lze zamítnout na hladině významnosti α .“ (Hendl, 2012, s. 190)

posteriorního rozdělení těchto parametrů zcela dostačující informací.“ (Hebák, 2012b, s. 87)

V této části se budeme nejprve zabývat tématem apriorního rozdělení a uvedeme typy rozdělení, která jsou nejčastěji používána v analýze sociálněvědních dat a poté popíšeme rozdělení důležitá pro bayesovský přístup. Dále už se budeme zabývat porovnáním bodových odhadů, intervalových odhadů, testování hypotéz a lineární regresi z hlediska frekvenční a bayesovské statistiky. V jednotlivých částech nejprve popíšeme krátce frekvenční přístup, poté bayesovský přístup a uvedeme konkrétní příklady analýzy na datech šetření CVVM Naše společnost – prosinec 2013. V závěru uvedeme, jaké jsou dostupné programy pro bayesovskou analýzu.

6.1 Apriorní rozdělení

Jak už bylo několikrát zmíněno, vstupní informací, kterou výzkumník do analýzy vkládá (před sběrem dat) je apriorní informace ve formě apriorního rozdělení. Apriorní rozdělení je založené na subjektivní pravděpodobnosti a je obvykle voleno na základě teorie, předchozích studií či výzkumníkovy zkušenosti (Gill, 2008, s. 7-8). Specifikování vhodného apriorního rozdělení parametru je nejvýraznějším aspektem bayesovské statistiky (Lynch, 2007, s. 54). Více než 100 let marného hledání způsobu jak ho obejít jasně demonstruje jeho důležitost (Gill, 2008, s. 44).

6.1.1 Pravděpodobnostní rozdělení

V příkladech na Bayesovu větu jsme uvažovali pravděpodobnost jednoho nebo více samostatných jevů. V analýze sociálněvědních dat však nepracujeme s jednotlivými pravděpodobnostmi, ale využíváme pravděpodobnostních funkcí, které obsahují určité parametry, jež určují tvar a měřítko křivky definované touto funkcí. Tato funkce se pro diskrétní proměnné označuje jako pravděpodobnostní funkce a pro spojité proměnné jako funkce hustoty pravděpodobnosti. Tyto funkce nám říkají, kde jsou ve výběrovém prostoru koncentrovány nejpravděpodobnější a nejméně pravděpodobné jevy (Lynch, 2007, s. 12). „Pravděpodobnostní rozdělení náhodné proměnné jsou pravděpodobnosti, s nimiž náhodná proměnná nabývá určité hodnoty nebo je obsažena v určitých intervalech hodnot.“ (Hendl, 2012, s. 132) Mohou se zdát složité, ale jsou to vlastně funkce, které popisují relativní frekvence výskytu pro konkrétní hodnoty náhodné proměnné (Lynch, 2007, s. 25). Pomocí apriorního rozdělení v bayesovské analýze vytváříme posteriorní rozdělení, která má také formu pravděpodobnostního rozdělení (Lynch, 2007, s. 50).

Abychom mohli určit rozdělení nějaké proměnné, musíme znát jeho parametry, kterými jsou očekávaná hodnota a rozptyl (neboli variance), který popisuje rozptýlenost hodnot náhodné proměnné od její očekávané hodnoty (Hendl, 2012, s. 134-135). Kromě rozptylu se někdy používá také směrodatná odchylka, což je druhá odmocnina z rozptylu.

6.1.2 Rozdělení v sociálních vědách

V sociálních vědách proměnné často sledují rozdělení, která mají vrchol ve středu a směrem k extrémním (krajním) hodnotám klesají, nebo mají vrchol na jednom konci rozdělení a směrem k druhému konci klesají (Lynch, 2007, s. 14). Mezi tato rozdělení patří normální, binomické, multinomické, Poissonovo a mnohorozměrné normální rozdělení.

Normální rozdělení je nejpoužívanějším rozdělením v sociálních vědách i statistice obecně. Jeho důležitost plyne z toho, že většina proměnných sleduje toto rozdělení (Lynch, 2007, s. 29). Díky centrálnímu limitnímu teorému lze mnoho rozdělení výběrových statistik aproximovat při větším rozsahu výběru normálním rozdělením, z čehož plyne i mimořádné postavení normálního rozdělení. Jeho základním tvrzením tedy je, že součet mnoha nezávislých libovolně rozdělených proměnných je s rostoucím počtem přibližně normálně rozdělen (Hendl, 2012, s. 152). Je základem například pro OSL regresi a mnoho dalších modelů.

Binomické rozdělení je běžné diskrétní rozdělení používané v sociálních vědách. Je základním rozdělením pro dichotomické proměnné. Toto rozdělení také aproximuje k normálnímu a jeho zvláštním případem je alternativní (Bernoulliho) rozdělení (Lynch, 2007, s. 25-26).

Multinomické rozdělení je zobecněním binomického rozdělení, ve kterém jsou více než dvě kategorie. V sociologii toto rozdělení sledují vícekategoriální nominální proměnné, což jsou proměnné s kvalitativně odlišnými kategoriemi (např. náboženské vyznání, sounáležitost k politické straně, atd.) (Lynch, 2007, s. 27-28).

Poissonovo rozdělení je diskrétní rozdělení, které poskytuje pravděpodobnosti pro daný počet úspěchů v nekonečném množství pokusů. Je často používáno k modelování počtu výsledků proměnných jako např. počet zatčení, počet dětí atd. Objevuje se zvláště u proměnných, u kterých se očekává nízký očekávaný počet, protože rozdělení takovýchto proměnných je často zešíkmené doprava s největším počtem hodnot blízko nuly (Lynch, 2007, s. 28).

Mnohorozměrné normální rozdělení je normální rozdělení pro více než jednu dimenzi. Je také poměrně často používaným rozdělením v sociálních vědách. Zvláště pak dvourozměrné normální rozdělení se používá pro vzájemné porovnání dvou výsledných proměnných, které jsou na sobě závislé (Lynch, 2007, s. 30-33).

6.1.3 Rozdělení v bayesovské statistice

Nutnost specifikovat apriorní rozdělení je označováno za nevýhodu bayesovské statistiky, ale ve skutečnosti to může být spíše výhoda, která nám poskytuje značnou flexibilitu – když nemáme žádnou informaci nebo máme jen málo informací o parametru, pak můžeme použít tzv. neinformativní apriorní rozdělení (*noninformative* nebo *diffuse*), který má jen velmi malý vliv na posteriorní rozdělení, neboť nefavorizuje žádnou z hypotéz a nechává v analýze dominovat data (Western, 1999, s. 11). O tomto přístupu se někdy mluví jako o objektivní bayesovské statistice (Lecoutre a Pointevineau, 2010). Při jeho použití jsou výsledky shodné s výsledky frekvenční statistiky (s metodou maximální věrohodnosti). Naopak pokud máme konkrétní informace o parametru, pak můžeme do analýzy vložit jasně specifikované subjektivní apriorní rozdělení (tzv. informativní rozdělení). To jsou v podstatě dva extrémy, mezi kterými můžeme volit. Zvolit můžeme samozřejmě také „něco mezi“ (Lynch, 2007, s. 55). V posledních dvaceti letech je používání subjektivních apriorních rozdělení na ústupu, naopak roste důraz na objektivní neinformativní apriorní rozdělení (Efron, 2005, s. 5).

V bayesovské statistice můžeme rozlišit tři obecné kategorie apriorních rozdělení: konjugované, informativní a neinformativní. Přesto, že je tato kategorizace praktická, je umělá, neboť hranice mezi nimi jsou neostře: konjugované formy jsou obvykle vysoce informativní a rozdíly mezi neinformativními a informativními jsou relativní (Gill, 2008, s. 136). Druhů apriorních rozdělení existuje ohromné množství (viz. Gill, 2008, kapitola 5), zde uvedu jen několik základních.

6.1.3.1 Konjugovaná apriorní rozdělení

Primární výhodou specifikace konjugovaných apriorních rozdělení je, že jsou matematicky jednodušší. Když apriorní rozdělení a věrohodnost sledují takové rozdělení, že posteriorní rozdělení má stejné rozdělení, pak jsou apriorní rozdělení a věrohodnost nazývány konjugovanými (Lynch, 2007, s. 57). Použití konjugovaného apriorního rozdělení je tedy způsob, jak zaručit, že posteriorní rozdělení bude mít jednoduše spočitatelnou formu (Gill, 2008, s. 138). Historicky byla v bayesovské

statistice konjugovaná apriorní rozdělení velmi důležitá, neboť usnadňovala analýzu. Před příchodem MCMC algoritmů bylo mnoho modelů příliš složitých, než aby je bylo možné určit bez nějakého triku jako je konjugace. To už dnes neplatí, konjugovaná apriorní rozdělení jsou však stále užitečným nástrojem (Gill, 2008, s. 138-144). Existují přehledné tabulky, které uvádí konjugovaná apriorní rozdělení a věrohodnosti (např. Gill, 2008, s. 142).

6.1.3.2 Informativní apriorní rozdělení

Informativní apriorní rozdělení jsou taková, která záměrně vkládají informaci, jež má výzkumník k dispozici. O velké většině sociálněvědních jevů máme před výzkumem nějaké informace, které díky bayesovské analýze můžeme zahrnout do statistické inference. Při používání informativních apriorních rozdělení je velmi důležité otevřeně doložit a obhájit používaná apriorní rozdělení, aby se tak zabránilo možnému nařčení ze záměrné manipulace analýzy k získání žádoucího výsledku. Vhodné je také při použití informativních apriorních rozdělení srovnat výsledky s analýzou používající neinformativní apriorní rozdělení (Gill, 2008, s. 156-157).

Z čeho vycházet při tvorbě informativních apriorních rozdělení? Obecně je tu spousta dostupných zdrojů, které mohou výzkumníka vést. Informativní apriorní rozdělení mohou vycházet z informací získaných z předchozích studií, publikovaných prací, od výzkumných institucí, či například pomocí rozhovorů s odborníky. Informační zdroje se samozřejmě mohou překrývat a v praxi se více zdrojů nemusí shodovat. Užitečnou strategií pak může být použití více soupeřících názorů, které můžeme pomocí výsledných posteriorních rozdělení porovnat. Informativní apriorní rozdělení můžeme dále rozlišit podle toho, z jakého zdroje informace čerpají. Můžeme rozlišit například tzv. *power priors*, které používají explicitně data z předchozích studií (Gill, 2008, s. 57). Dalšími jsou například tzv. *elicited priors*, které záměrně nevychází z matematických vlastností, ale znalostí o předmětu výzkumu poskytnutých experty bez znalosti statistiky (Gill, 2008, s. 159).

6.1.3.3 Neinformativní apriorní rozdělení

Neinformativní apriorní rozdělení jsou z pohledu bayesovské statistiky velmi užitečná. Neinformativní rozdělení přes svůj název není zcela neinformativním, neboť deklarováním neznalosti o parametru, je také vložení určité informace (Lynch, 2007, s. 55). Neinformativní apriorní rozdělení je takové, které poskytuje jen velmi málo explanační síly do analýzy o neznámém parametru. Vhodná jsou jednak pro situace, kdy

o studovaném jevu nemáme žádné předchozí informace, které bychom mohli použít pro vytvoření informativního apriorního rozdělení. Dále jsou vhodná také v případech, kdy se záměrně snažíme vložit do analýzy málo subjektivní informace, čímž lze částečně zmírnit frekvenční kritiku subjektivity (Gill, 2008, s. 144). Neinformativní apriorní rozdělení se využívají v tzv. objektivní bayesovské statistice.

Nejdůležitějším neinformativním apriorním rozdělením je rovnoměrné (*uniform*) rozdělení. Rovnoměrné rozdělení se samo o sobě v sociologických datech příliš často neobjevuje, neboť velmi málo fenoménů v sociálních vědách toto rozdělení sleduje (Lynch, 2007, s. 14). Rovnoměrné rozdělení modeluje situaci, ve které se žádné hodnotě nedává přednost (Hebák, 2012b, s. 84).

6.2 Bodový odhad

6.2.1 Frekvenční přístup k bodovým odhadům

Bez základní znalosti frekvenční teorie bodových odhadů není možné se pokoušet o její kritické hodnocení či srovnání s bayesovským. Jak už bylo uvedeno, frekvenční statistika používá k úsudkům o neznámých parametrech problematického výběrového rozdělení. Pojem výběrové rozdělení stojí v sociologii na nerealistickém předpokladu „... *nezávislých všech možných náhodných výběrů pořizovaných za zcela stejných podmínek.*“ (Hebák, 2012b, s. 78; zvýraznění v originále) Všechny uvažované pravděpodobnosti používané při frekvenčních statistických úsudcích proto „... nemají nic společného s daty z pořizovaného náhodného výběru a jsou stanoveny předem, tedy dříve než výběr byl proveden. Tím však dochází k tomu, že kvalita použitelných *statistik* je posuzována na základě nerealistické abstrakce a teoretické (částečně hypotetické) úvah.“ (Hebák, 2012b, s. 78; zvýraznění v originále)

6.2.2 Bayesovský přístup k bodovým odhadům

Statistické úsudky v bayesovském přístupu vycházejí ze spočítané posteriorní informace, která je proporcionální součinu apriorní informace a výběrové informace. Posteriorní rozdělení vyjadřuje náš aktualizovaný a modifikovaný (po získání a zahrnutí daty) postoj k hodnotě parametru (Hebák, 2012b, s. 79-80).

Klasické statistické úsudky vycházejí z výběrového rozdělení v situaci před provedením výběru, a proto pravděpodobnosti a z nich plynoucí úsudky „... lze vztahovat jen k této situaci a nemají nic společného s konkrétními daty ze zamýšleného nebo už provedeného výběru.“ (Hebák, 2012b, s. 82) „Ve srovnání s tím se

bayesovským bodovým odhadem rozumí **konkrétní hodnota** charakteristiky posteriorního rozdělení, které je s pomocí **apriorního rozdělení** $g(\theta)$ vypočítaná na základě dat z provedeného **výběru**.“ (Hebák, 2012b, s. 82; zvýraznění v originále) Bolstad (2007, s. 169) označuje frekvenční odhady za *pre-data* úsudky, neboť jsou založené na výběrovém rozdělení. Naopak bayesovské odhady označuje za *post-data* úsudky, neboť vycházejí z napozorovaných dat.

6.3 Intervalové odhady

6.3.1 Frekvenční intervaly spolehlivosti

Konstrukce frekvenčního intervalu spolehlivosti je opět založená na výběrovém rozdělení použité statistiky, z čehož vyplývá, že stejně jako u bodových odhadů, pravděpodobnosti s tím související jsou určeny „... *dříve, než máme data*, a tedy nezávisí na konkrétním posuzovaném výběru.“ (Hebák, 2012c, s. 76; zvýraznění v originále)

Stejně potíže s použitím pravděpodobnosti po tom, co získáme výběrová data, nalezneme i v interpretaci intervalů spolehlivosti. Teorie říká, že z dlouhodobého hlediska s daty z mnoha opakovaných výběrů (viz. výběrové rozdělení), a tudíž mnoha intervaly spolehlivosti, určitá proporce těchto intervalů spolehlivosti bude obsahovat skutečnou hodnotu parametru, zatímco ostatní nebudou. Teorie tedy předpovídá, co se stane z dlouhodobého hlediska, před tím než jsou nasbírána data. Potíž je v tom, co udělat s pravděpodobnostmi po tom, co jsou data z našeho jednoho výběru známá. Jediný interval spolehlivosti z našeho (jediného) výběru buď obsahuje, nebo neobsahuje skutečný parametr. My ale nevíme, jestli náš interval patří do velké skupiny intervalů (95 ze 100), které parametr obsahují, nebo do malé skupiny intervalů (5 ze 100), které parametr neobsahují. Jediné, co můžeme udělat, je doufat, že náš jediný známý interval patří do první skupiny (Iversen, 1984, s. 11). Frekvenční interval spolehlivosti je tak do jisté míry spíše hypotetickou úvahou.

Běžně jsou intervaly spolehlivosti interpretovány pravděpodobnostně (tedy tak, jako by to byly intervaly vycházející z bayesovského přístupu), že daný parametr se spolehlivostí 95 % leží uvnitř tohoto intervalu. Teoretický parametr však nepředstavuje náhodnou proměnnou, a proto také nelze mluvit o pravděpodobnostech jeho hodnot (Hendl, 2012, s. 177). „Správná interpretace klasického intervalu spolehlivosti proto je, že $(1 - \alpha)100\%$ náhodných intervalů ze všech možných, vypočítaných tímto způsobem, *bude* obsahovat skutečnou hodnotu neznámého parametru π .“ (Hebák, 2012c, s. 76;

zvýraznění v originále) Podle bayesiánů jsou pravděpodobnostní úsudky založené na takto pojatých intervalech spolehlivosti chybné a zavádějící.

6.3.2 **Bayesovské kredibilní intervaly**

Bayesovská alternativa intervalů spolehlivosti vychází z posteriorního rozdělení a je označována za kredibilní interval, resp. interval kredibility. Bayesovský přístup sumarizuje informaci získanou z konkrétních dat našeho výběru. Pravděpodobnostní povaha tohoto intervalu umožňuje jeho jednoduchou interpretaci. Při tomto pojetí se na rozdíl od frekvenčního přístupu „... nemusíme zajímat o skutečnosti, které mohly nastat, ale nenastaly.“ (Hebák, 2012c, s. 76).

Hebák shrnuje, že „... chybně používaná interpretace klasického intervalu spolehlivosti je naopak správnou interpretací bayesovského kredibilního intervalu, jehož konstrukce výhradně vychází z posteriorního rozdělení parametru θ .“ (Hebák, 2012c, s. 77) Jako kredibilní interval se nejčastěji volí tzv. *highest probability density region* (HPD region). HPD region nemusí tvořit pouze jeden interval, ale podle tvaru rozdělení může být tvořen několika intervaly (Jackman, 2009, s. 27-28).

Bayesovské kredibilní intervaly jsou většinou kratší (užší) než klasické intervaly spolehlivosti, a to i při vágní apriorní informaci. Rozdíly mezi nimi jsou pak tím větší, čím lepší je apriorní informace (Hebák, 2012c, s. 77).

6.3.3 **Příklad: Odhady průměrného věku**

V této části budeme porovnávat bodový a intervalový odhad průměrného věku respondenta na datech z výzkumu CVVM Naše společnost 2013 – prosinec. Tento výzkum jsme zvolili proto, že zkoumané jevy (proměnné) se v dotaznících každý rok opakují a je tedy možné využít předchozí znalosti o daném jevu k vytvoření subjektivního (informativního) apriorního rozdělení. Předchozí znalosti o průměrném věku bychom samozřejmě mohli získat i z jiných zdrojů¹⁹, v tomto případě však můžeme využít toho, že oba výzkumy mají stejnou metodologii. Bodový odhad porovnáme pro tři různá apriorní rozdělení (objektivní, subjektivní a záměrně chybně

¹⁹ Apriorní rozdělení průměrného věku respondenta můžeme samozřejmě také určit na základě teoretické úvahy. Například víme, že respondenti jsou ve věku od 15 do přibližně 90 let. Myslíme si ale, že to bude o něco méně, než je střed tohoto intervalu (52,5 let), řekněme 48 let. Směrodatnou odchylku bychom nastavili např. na 15 (tzn., že věříme, že průměr leží někde mezi 33 a 63 lety).

zvolené) a pro tři různé velké velikosti výběru – na původních datech $N = 987$ a na datech vytvořených náhodným výběrem (10 % a 5 % z původních dat).

Volba objektivního apriorního rozdělení byla ponechána jako výchozí nastavení programu BugsXLA, který defaultně vytváří neinformativní apriorní rozdělení (tzv. *flat priors*). Jde o apriorní rozdělení nastavená tak, aby odpovídala odhadům vytvořeným pomocí metody maximální věrohodnosti, která se používá ve frekvenční statistice. U proměnných je očekávaná hodnota rozdělení proměnné nastavena jako průměr dané proměnné a směrodatná odchylka je nastavena jako stonásobek směrodatné odchylky dané proměnné. Jako subjektivní apriorní rozdělení bylo zvoleno rozdělení, které jsme získali na základě dat z předchozího výzkumu CVVM Naše společnost 2012 – prosinec²⁰. Proměnná „věk respondenta“ v průzkumu v roce 2012 měla přibližně normální rozdělení se střední hodnotou 44,38 a rozptylem 298,147 (a směrodatnou odchylkou 17,267), což lze také zapsat jako $\sim N(44,38, 298,147)$. Jako záměrně chybné apriorní rozdělení jsme zvolili rozdělení, jehož střední hodnota i rozptyl je 100 (a směrodatná odchylka je 10), neboli $\sim N(100, 100)$. Chybné rozdělení tedy do analýzy vkládá informaci, že proměnná věk má normální rozdělení s průměrem 100 let a směrodatnou odchylkou 10 let.

Nejdříve pro ilustraci ukažme, jak vypadá výstup odhadu pomocí subjektivní bayesovské analýzy:

²⁰ Použili jsme tedy posteriorní rozdělení již provedené analýzy (na datech z roku 2012) jako apriorní rozdělení nové analýzy (na datech z roku 2013).

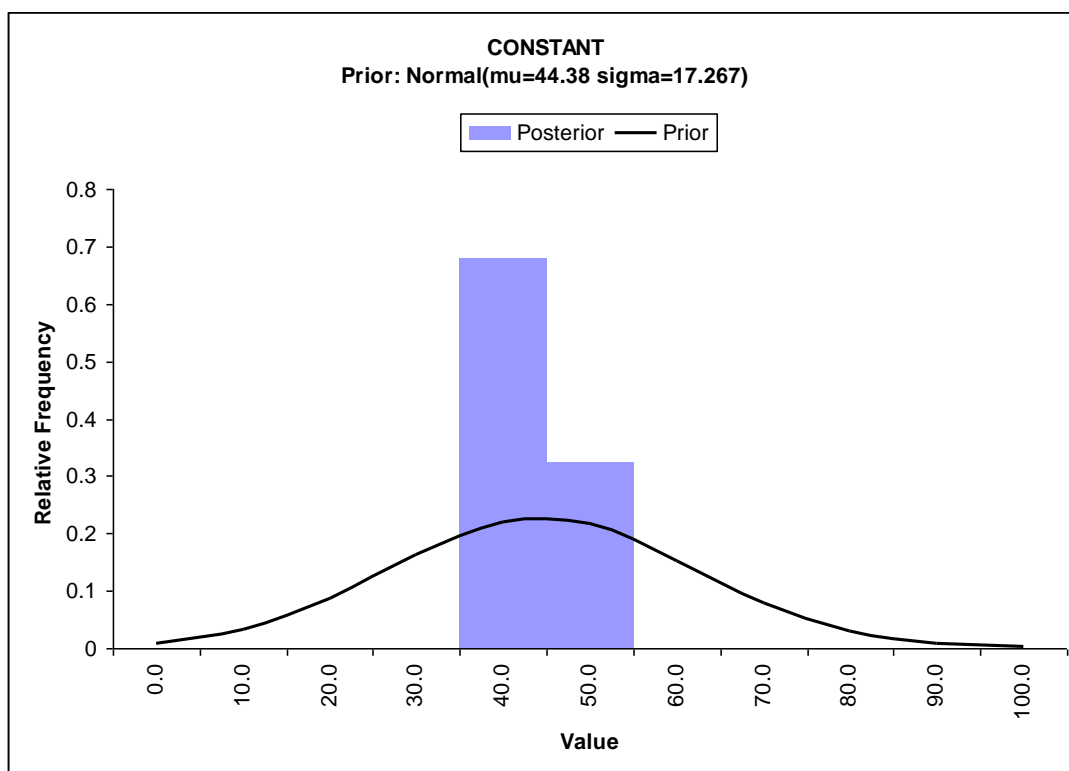
Tabulka 1. Odhad průměrného věku pomocí subjektivního apriorního rozdělení v programu BugsXLA (ukázka výstupu).

| | | Label | Mean | St.Dev. | 2.5% | Median | 97.5% |
|--|----------------------------|--------------|-------------|----------------|-------------|---------------|--------------|
| | | CONSTANT | 46.91 | 1.8460 | 43.28 | 46.8800 | 50.59 |
| | | SD(residual) | 18.34 | 1.3120 | 16.02 | 18.2600 | 21.17 |
| | | | | | | | |
| Model | [Desetina!\$A\$1:\$D\$101] | | | | | | |
| Distribution | Normal | | | | | | |
| Link | Identity | | | | | | |
| Response | IDE_2 | | | | | | |
| | | | | | | | |
| Priors | | | | | | | |
| CONSTANT | N(mu=44.4, sigma=17.3) | | | | | | |
| V(residual) | Inv-Gamma(0.001, 0.001) | | | | | | |
| | | | | | | | |
| WinBUGS MCMC Settings | | | | | | | |
| Burn-In: 5000 Samples: 1000 (Thin:1; Chains:3) | | | | | | | |
| Run took 13 seconds (Auto Quit OFF) | | | | | | | |
| BugsXLA (Beta 5.0) 2011.Apr.17.(00.00) | | | | | | | |

Zdroj: CVVM Naše společnost – prosinec 2013, N = 100. Vlastní výpočty.

Výsledek analýzy nejdříve shrnuje bodové a intervalové odhady proměnné (CONSTANT), v základním nastavení program nabízí informace o směrodatné odchylce, odhad průměru, mediánu a kvantily 95 % intervalu kredibility. V části „Model“ vidíme informace názvu souboru (Desetina znamená, že jde o data představující výběr 10 % hodnot původních dat) a typu rozdělení (normální). V části nazvané „Priors“ vidíme vstupní apriorní rozdělení proměnné, které se skládá z průměru a směrodatné odchylky. V části „WinBUGS MCMC Settings“ vidíme nastavení MCMC algoritmu a další informace. Apriorní a posteriorní rozdělení proměnné můžeme zobrazit najednou v přehledném grafu:

Graf 1. Ukázka grafu zobrazujícího apriorní a posteriorní rozdělení parametru.



Zdroj: CVVM Naše společnost – prosinec 2013, N = 100. Vlastní výpočty.

Výsledky analýz pro jednotlivé výběrové soubory vidíme v následujících tabulkách:

Tabulka 2. Odhad průměrného věku respondenta pro N = 987.

| Analýza | Bodový odhad | Interval 2,5 % | Interval 97,5 % |
|------------------------|--------------|----------------|-----------------|
| Frekvenčně | 44,77 | 43,68 | 45,85 |
| Bayesovsky objektivně | 44,73 | 43,61 | 45,88 |
| Bayesovsky subjektivně | 44,73 | 43,61 | 45,88 |
| Bayesovsky chybně | 44,89 | 43,79 | 46,05 |

Zdroj: CVVM Naše společnost – prosinec 2013. N = 987. Vlastní výpočty.

Tabulka 3. Odhad průměrného věku respondenta pro N = 100.

| Analýza | Bodový odhad | Interval 2,5 % | Interval 97,5 % |
|------------------------|--------------|----------------|-----------------|
| Frekvenčně | 46,93 | 43,32 | 50,54 |
| Bayesovsky objektivně | 46,94 | 43,28 | 50,65 |
| Bayesovsky subjektivně | 46,91 | 43,28 | 50,59 |
| Bayesovsky chybně | 48,69 | 45,03 | 52,42 |

Zdroj: CVVM Naše společnost – prosinec 2013. N = 100. Vlastní výpočty.

Tabulka 4. Odhad průměrného věku respondenta pro N = 53.

| Analýza | Bodový odhad | Interval 2,5 % | Interval 97,5 % |
|-------------------------------|---------------------|-----------------------|------------------------|
| Frekvenčně | 46,81 | 41,75 | 51,87 |
| Bayesovsky objektivně | 46,87 | 41,84 | 52,06 |
| Bayesovsky subjektivně | 46,82 | 41,84 | 51,97 |
| Bayesovsky chybně | 50,29 | 45,26 | 55,69 |

Zdroj: CVVM Naše společnost – prosinec 2013. N = 53. Vlastní výpočty.

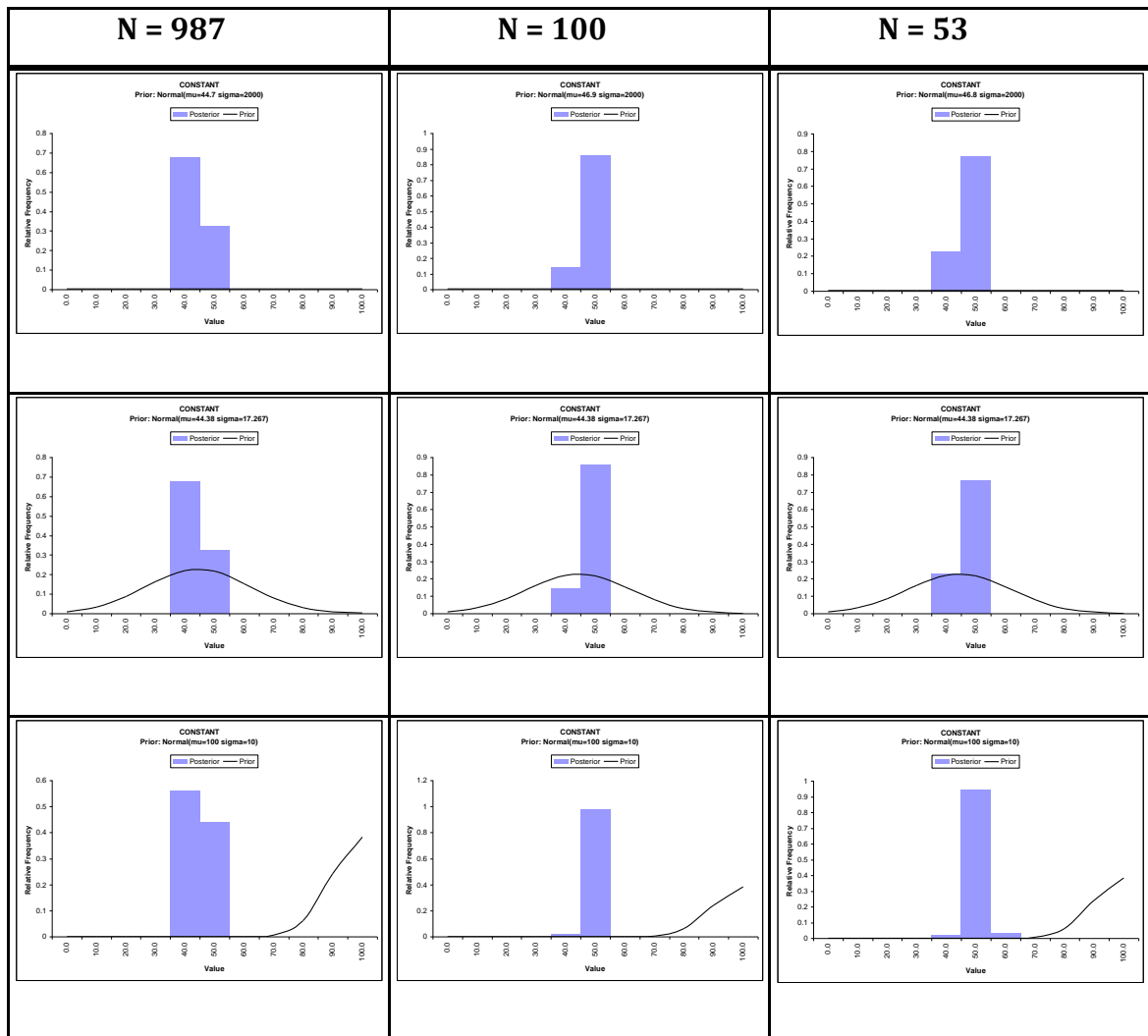
Z analýzy vyplývá, že bodové a intervalové odhady prováděné pomocí frekvenční a objektivní bayesovské analýzy se liší jen minimálně – v našem příkladu jen v řádu setin. Objektivní bayesovský přístup a subjektivní bayesovský přístup se v případě původních dat shodují, u 10% výběru se liší o tři desetiny a u 5% výběru se liší o pět setin. U záměrně chybné bayesovské analýzy se výsledky od objektivní bayesovské analýzy liší, tak jak jsme předpokládali. Platí, že čím větší je velikost výběru, tím větší vliv mají na apriorní rozdělení data a naopak. To je demonstrováno na příkladu záměrně chybné bayesovské analýzy – průměrný věk se vzdaluje se zmenšujícím výběrem od průměrné hodnoty frekvenčnímu a tudíž i objektivnímu bayesovskému odhadu.

Co se týče interpretace odhadů, u frekvenční analýzy můžeme říci, že bodový odhad, tedy skutečnost, že průměrný věk respondenta je 44,77 let, není spočítán na základě výběrových dat, ale je dán předem. Interval spolehlivosti správně interpretujeme tak, že 95 % náhodných intervalů ze všech možných, vypočítaných tímto způsobem, bude obsahovat skutečnou hodnotu průměrného věku respondenta. U bayesovské analýzy je interpretace odhadů podstatně jednodušší. Bodový odhad průměru, tedy hodnota například 44,73, představuje skutečně průměrný věk spočítaný na základě výběrových dat. Intervalový odhad, neboli kredibilní interval interpretujeme tak, že průměrný věk respondenta se spolehlivostí 95 % leží uvnitř tohoto intervalu.

Na následujících grafech jsou dobře patrné rozdíly mezi apriorním rozdělením (zobrazeno jako křivka) a posteriorním rozdělením (zobrazeno jako histogram) a jejich tvarem. Sloupce v tabulce představují jednotlivé velikosti výběrů. První řádek představuje neinformativní (objektivní) apriorní rozdělení, které je ploché (tzv. *flat*) a v podstatě kopíruje osu x . V druhé řádce je subjektivní apriorní rozdělení zvolené na

základě předchozí informace o průměrném věku respondentů. Třetí řádek představuje záměrně špatnou apriorní informaci, u které jak je vidět, že má mírný vliv na tvar posteriorního rozdělení („posouvá“ jeho hodnoty směrem doprava).

Tabulka 5. Grafy apriorních a posteriorních rozdělení průměrného věku respondenta pro různě velké datové soubory.



Zdroj: CVVM Naše společnost – prosinec 2013. N = 987, N = 100, N = 53. Vlastní výpočty.

6.4 Testování hypotéz

6.4.1 Klasický přístup k testování hypotéz

Klasický přístup k testování hypotéz je obecně tím nejvíce kritizovaným ze všech statistických úsudků (a to nikoliv jen bayesiány) (Hebák, 2012c, s. 78). Testy

významnosti podléhají kritice už od jejich vzniku a tato kritika stále trvá. V sociálních vědách se zvedla zejména v 60. letech minulého století. Test významnosti nám řekne, zda výsledek je signifikantní, nebo není. Vzhledem k tomu, že testy významnosti jsou používány i při analýze dat, která nejsou experimentální, vede jejich používání často ke zneužívání ve výběru a interpretaci výsledků. Navíc ve vědeckých časopisech jsou většinou publikovány pouze signifikantní výsledky, ty nesignifikantní výsledky jsou zcela přecházeny (Lecoutre a Pointevineau, 2010). Námitky stoupců bayesovského přístupu vůči testování hypotéz frekvenčním způsobem se zakládají opět na problematickém výběrovém rozdělení, tak jako u bodových a intervalových odhadů. Bayesiáni kritizují i možnost volby hladiny významnosti, která ve svém důsledku je jedinou rozhodovací dělicí čarou (Hebák, 2012c, s. 80).

Postup frekvenčního testování hypotéz²¹ probíhá ve čtyřech krocích. V prvním kroku formulujeme výzkumnou otázku ve formě nulové hypotézy, která tvrdí, že nalezený rozdíl lze přičíst přirozené variabilitě dat, a alternativní statistické hypotézy, která tvrdí, že nulová hypotéza neplatí. V druhém kroku volíme přijatelnou úroveň chyby rozhodování pomocí hladiny významnosti alfa, která se volí obvykle 0,05. V třetím kroku se vypočítá testovací statistika. V konečném čtvrtém kroku obvykle zhodnotíme hodnotu významnosti p , která říká „... jaká je pravděpodobnost, že získáme vypočítanou hodnotu nebo ještě neobvyklejší hodnotu testovací statistiky, jestliže nulová hypotéza platí.“ (Hendl, 2012, s. 183) V případě, že je p -hodnota menší nebo rovná hladině významnosti alfa, pak zamítáme nulovou hypotézu. Jak Hendl podotýká, nezamítnutí nulové hypotézy neznamená její důkaz, ale znamená spíše, že nemáme dostatek evidence pro její zamítnutí (Hendl, 2012, s. 182-183).

Iversen (1984) Popisuje příklad testování hypotézy, zda je v populaci 50 % žen a 50 % mužů. Nulová hypotéza zní, že 50 % náhodně vybraných osob budou ženy. Budeme náhodně vybírat vzorek deseti lidí a uděláme dvoustranný test. Pokud je nulová hypotéza pravdivá, pak můžeme použít binomické rozdělení k nalezení pravděpodobností, že výběr obsahuje 0 žen (pravděpodobnost 1/1024), 1 ženu (10/1024), 9 žen (10/1024) a 10 žen (1/1024). Pokud zamítneme nulovou hypotézu pro 0, 1, 9 nebo 10 žen, pak se hodnota signifikance tohoto testu bude rovnat součtu

²¹ Ve skutečnosti existují dva přístupy k testování hypotéz – Fisherův přístup a jeho obdoba podle Neymana a Pearsona. Je zajímavé, že přesto dnes přístupy Neymana-Pearsona a Fishera splývají.

pravděpodobností těchto případů, který je $22/1024 = 0,02$. Nyní provedeme výběr, ve kterém je 9 žen. Protože výsledek spadá do oblasti zamítnutí, pak můžeme říci, že zamítáme nulovou hypotézu s úrovní signifikance 0,02. Úroveň signifikance sestává částečně z pravděpodobnosti napozorovaných dat, ale obsahuje také pravděpodobnosti pro 0, 1 a 10 žen, což jsou data, která nikdy nenastala. V klasické statistice jsme tedy ve zvláštní situaci, kdy pravděpodobnosti různých dat, která nenastala, slouží jako důkaz proti nulové hypotéze. Lze tedy říci, že „... zamítnutí nulové hypotézy je založené na pravděpodobnosti z dat, která nebyla napozorována.“ (Lecoutre a Pointevineau, 2010)

Důvod tohoto úskalí tkví opět v tom, že teorie frekvenční inference je založená na pravděpodobnostech jako dlouhodobých relativních četnostech. Hodnota signifikance nám říká, co se stane z dlouhodobého hlediska, když vybereme velké množství výběrů. V sociologii však obvykle nemáme mnoho výběrů, máme naopak pouze jeden výběr (Iversen, 1984, s. 10).

Po uskutečnění výběru se stane obtížnějším interpretovat hodnotu signifikance spojenou s jedním výběrem. Problematika interpretace a chápání, co je signifikance, je dokonce studována a ukazuje se, že se správnou interpretací mají problém nejen studenti, ale také jejich vyučující a profesori (Haller a Krauss, 2002). Ve skutečnosti je p -hodnota „... pravděpodobnost (jakoby počítaná před provedením výběru za předpokladu pravdivosti H) získat již získaný výběrový výsledek či výsledek ještě nepříznivější testované hypotéze H .“ (Hebák et al., 2007, s. 60-61; zvýraznění v originále) Jinými slovy „... p -hodnoty říkají, jak často můžeme při provedení všech možných výběrů a při platnosti hypotézy H_0 získat vypočítanou hodnotu testového kritéria anebo hodnoty, které jsou pro testovanou hypotézu ještě více nepříznivé. Taková interpretace působí velice neohrabaně až podezřele, takže je zcela pochopitelná snaha p -hodnoty jednoduše označit za hladinu významnosti, se kterou jsou porovnávány, což také (označením *sig* ve zkráceném smyslu *významnost*) mnohé výstupy statistických paketů obsahují.“ (Hebák, 2012c, s. 80-81; zvýraznění v originále)

6.4.2 Bayesovský přístup k testování hypotéz

Bayesovské testování hypotéz je často méně formální než u nebayesovských variant. Navíc nejběžnější procedurou pro sumarizování výsledků v sociálněvědním výzkumu je jednoduché popsání posteriorního rozdělení spíše než aplikace rigidního rozhodovacího procesu (Gill, 2008, s. 229). Klasické testování hypotéz nám dokáže pomoci rozhodnout se mezi dvěma volbami – nulovou a alternativní hypotézou (Gill,

2008, s. 232) Bayesiáni se o testování hypotéz zajímají obecně méně, neboť posteriorní rozdělení dokáže poskytnout o neznámém parametru vše potřebné. Někdy mohou být však testy významnosti užitečné, a protože jsou v bayesovském přístupu vyjádřeny pravděpodobnostně, je snadné pomocí nich srovnat konkurující si tvrzení (Hebák, 2012c, s. 80).

Bayesovská verze p -hodnoty může být spočítána z posteriorního rozdělení (Gill, 2008, s. 233-234). Přestože je na výpočet obtížnější, než klasická p -hodnota, je mnohem užitečnější a její interpretace je také mnohem jednodušší – pravděpodobnost, že nulová hypotéza platí, za předpokladu dat a modelu (Lecoutre a Pointevineau, 2010). P -hodnotu obecně vypočítáme jako podmíněnou pravděpodobnost nulové hypotézy za předpokladu napozorovaných dat. Výsledek interpretujeme tak, že je-li posteriorní pravděpodobnost menší než zvolená hladina významnosti, pak nulová hypotéza platí (a naopak) (Hebák, 2012c, s. 80).

Nejběžnější frekvenční forma testování hypotéz s nulovou hypotézou $H_0: P = 0$ je v bayesovském přístupu nerealizovatelná, pokud P je spojitý parametr (u spojitých proměnných je pravděpodobnost jedné konkrétní hodnoty rovna nule) (Jackman, 2009, s. 31). Hypotézy je však možné zvolit jako $H_0: P \leq 0$ a $H_1: P > 0$ (Hebák, 2012c, s. 81). Případně je možné definovat nulovou hypotézu jako malý interval okolo nulové hypotézy $H_0: -\varepsilon \leq P \leq \varepsilon$. Může se to zdát na jednu stranu jako nevýhoda, ale u frekvenčního testování hypotéz také prakticky nikdo nevěří, že nějaký testovaný parametr se rovná přesně nule (Gill, 2008, s. 238). Další možností je postupovat pomocí bayesovského dvoustranného kredibilního intervalu. V takovém případě se testovaná hypotéza zamítá v případě, že se testovaná hodnota v tomto intervalu nenachází. Nezamítá se naopak v případě, že se testovaná hodnota v tomto intervalu nachází (Hebák, 2012c, s. 81).

6.4.3 Příklad: Dvouvýběrový t-test

V této části jako praktickou ukázkou testování hypotéz provedeme dvouvýběrový t-test. Z důvodu naprosto odlišného vytváření nulové a alternativní hypotézy nelze frekvenční a bayesovský přístup zcela korektně porovnat (Gill, 2008, s. 239). Posteriorní pravděpodobnost hypotézy je smysluplná jen pro bayesovský rámec, ve frekvenčním přístupu nic takového neexistuje (Jackman, 2009, s. 33).

Cílem této ukázky není variovat apriorní rozdělení proměnných, ale předvést výstup testu v programu BugsXLA, interpretovat jeho výsledek a porovnat jej

s frekvenčním výstupem. V BugsXLA nenajdeme samostatnou proceduru, pomocí které bychom mohli t-testy provádět. Abychom ukázali oba možné výsledky testování hypotéz, předvedeme dva modely. V prvním modelu se průměry mezi skupinami budou lišit, v druhém se naopak lišit nebudou.

V prvním modelu se ptáme, zda se liší průměrný osobní měsíční příjem u mužů a žen. Nulová hypotéza zní, že se průměry neliší, zatímco alternativní zní, že se průměry liší. Jako výsledek dvouvýběrového t-testu v SPSS získáme následující tabulky:

Tabulka 6. Popisné statistiky t-testu v SPSS.

| | Pohlaví | N | Průměr |
|--------|---------|-----|--------|
| Příjem | Muž | 375 | 17560 |
| | Žena | 403 | 12910 |

Zdroj: CVVM Naše společnost – prosinec 2013. N = 778. Vlastní výpočty.

Tabulka 7. Výsledky t-testu v SPSS.

| | | T-test shody průměrů | | | | |
|--------|-------------------|-----------------------------|------|----------------|-----------------------------|-------|
| | | Levenův test shody rozptylů | | Sig.(2-tailed) | 95% Intervaly spolehlivosti | |
| | | F | Sig. | | Spodní | Horní |
| Příjem | Shodné rozptyly | 18.764 | .000 | .000 | 3212 | 6091 |
| | Neshodné rozptyly | | | .000 | 3177 | 6126 |

Zdroj: CVVM Naše společnost – prosinec 2013. N = 778. Vlastní výpočty.

V první tabulce vidíme průměrný plat mužů (17561 Kč) a žen (12910 Kč) a početní zastoupení obou skupin. V druhé tabulce vidíme Levenův test shody rozptylů, podle kterého se rozptyly statisticky významně liší. Podle toho se u t-testu shody průměrů díváme na druhý řádek, který shodu rozptylů nepředpokládá. Signifikance je menší než stanovená hladina významnosti 0,05, můžeme tedy říci, že průměry mužů a žen se od sebe statisticky významně liší.

Nyní provedeme obdobu tohoto testu bayesovsky a dostáváme:

Tabulka 8. Popisné statistiky obdoby t-testu v BugsXLA.

| Odhady průměru pro příjem | |
|---------------------------|--------|
| Pohlaví | Průměr |
| Muž | 17560 |
| Žena | 12920 |

Zdroj: CVVM Naše společnost – prosinec 2013. N = 778. Vlastní výpočty.

Tabulka 9. Výsledky obdoby t-testu v BugsXLA.

| Test závislosti příjmu na pohlaví | | |
|-----------------------------------|----------------------|--------|
| | Kredibilní intervaly | |
| | 2,5 % | 97,5 % |
| Příjem vs. pohlaví | -6127 | -3197 |

Zdroj: CVVM Naše společnost – prosinec 2013. N = 778. Vlastní výpočty.

Bayesovský výstup t-testu je poměrně podobný. V první tabulce opět vidíme průměrné platy dle pohlaví. V bayesovské statistice nic jako p -hodnota neexistuje, není tedy součástí výstupu. Součástí výstupu nicméně jsou kvantily 95% intervalu kredibility. Vzhledem k tomu, že intervaly kredibility neobsahují nulu, přijímáme alternativní hypotézu, že průměry příjmů u mužů a žen se liší.

V druhém příkladě se ptáme, zda se liší průměrný příjem u lidí se středoškolským vzděláním s maturitou a ostatními (dichotomická proměnná).

Tabulka 10. Popisné statistiky t-testu v SPSS.

| | Pohlaví | N | Průměr |
|--------|---------|-----|--------|
| Příjem | SSM | 216 | 16039 |
| | Ostatní | 562 | 14811 |

Zdroj: CVVM Naše společnost – prosinec 2013. N = 778. Vlastní výpočty.

Tabulka 11. Výsledky t-testu v SPSS.

| | | T-test shody průměrů | | | | |
|----------|-------------------|-----------------------------|-------|----------------|-----------------------------|-------|
| | | Levenův test shody rozptylů | | Sig.(2-tailed) | 95% Intervaly spolehlivosti | |
| | | F | Sig. | | Spodní | Horní |
| Vzdělání | Shodné rozptyly | 0,230 | 0,632 | 0,144 | -2872 | 418 |
| | Neshodné rozptyly | | | 0,106 | -2718 | 264 |

Zdroj: CVVM Naše společnost – prosinec 2013. N = 778. Vlastní výpočty.

Na základě frekvenčního dvouvýběrového t-testu zjišťujeme, že průměrný příjem respondentů s maturitou je 16039 Kč a ostatních je 14811 Kč. Levenův test shody rozptylů říká, že rozptyly se statisticky významně neliší. Signifikance příslušného t-testu je 0,144, což je více než stanovená hladina významnosti 0,05. Zjišťujeme tedy, že jsme ve zmíněných dvou skupinách neprokázali odlišnost mezi průměrnými příjmy.

Nyní opět provedeme obdobnou analýzu bayesovsky a dostáváme:

Tabulka 12. Popisné statistiky obdoby t-testu v BugsXLA.

| Odhady průměru pro příjem | |
|---------------------------|--------|
| Vzdělání | Průměr |
| SSM | 16040 |
| Ostatní | 14820 |

Zdroj: CVVM Naše společnost – prosinec 2013. N = 778. Vlastní výpočty.

Tabulka 13. Výsledky obdoby t-testu v BugsXLA.

| Test závislosti příjmu na vzdělání | | |
|------------------------------------|----------------------|--------|
| | Kredibilní intervaly | |
| | 2,5 % | 97,5 % |
| Příjem vs. vzdělání | -472 | 2873 |

Zdroj: CVVM Naše společnost – prosinec 2013. N = 778. Vlastní výpočty.

Opět získáváme tabulky pro průměry porovnávaných skupin. Na základě kredibilních intervalů, které obsahují nulu, můžeme říci, že se průměry v porovnávaných skupinách neliší.

Na základě příkladů jsme ukázali, že výsledky frekvenčního a bayesovského dvouvýběrového t-testu se z hlediska postupu odlišují, ale z hlediska jejich výsledků jsou shodné²². Frekvenční testování hypotéz postupuje (subjektivním nebo spíše konvenčním) stanovením hladiny významnosti alfa (tj. pravděpodobnosti, že zamítneme nulovou hypotézu, ačkoli ona platí). Na základě Levenova testu shody rozptylů se nejdříve rozhodujeme, který z výsledků t-testu je vhodný pro naše data (tj. zda mají skupiny shodné nebo odlišné rozptyly). Podle hodnoty signifikance poté rozhodujeme o tom, zda „... data podporují, nebo nepodporují rozhodnutí o zamítnutí platnosti nulové hypotézy.“ (Hendl, 2012, s. 190) Nulovou hypotézu zamítneme v případě, že je signifikance (neboli p -hodnota) menší než zvolená hodnota alfa. Nezamítneme ji v případě, že je signifikance větší než alfa. Nezamítnutí nulové hypotézy ovšem neznamená její přijetí, ale znamená, že nemáme dostatek evidence k jejímu zamítnutí. V případě bayesovské statistiky neexistuje nic jako hodnota alfa nebo signifikance. Rozhodnutí o přijetí nebo zamítnutí nulové hypotézy se rozhodujeme (v případě programu BugsXLA) pomocí toho, zda kredibilní intervaly obsahují nebo neobsahují nulu²³. Interpretace výsledků je v případě bayesovské alternativy dvouvýběrového t-testu jednodušší. Pokud kredibilní intervaly nulu neobsahují, pak můžeme s 95 % pravděpodobností říci, že se průměry mezi skupinami liší (zamítáme nulovou hypotézu). Pokud kredibilní intervaly nulu obsahují, pak zjišťujeme, že s 95 % pravděpodobností se průměry v porovnávaných skupinách neliší (přijímáme nulovou hypotézu).

6.5 Regresní analýza

Regresní analýza zkoumá statistickou závislost pomocí modelu, který obsahuje jednu závislou proměnnou a jednu nebo více nezávislých proměnných. V sociologii je nejběžněji používaná vícenásobná lineární regresní analýza, na jejímž příkladě opět porovnáme bayesovský a frekvenční přístup.

²² V bayesovské analýze jsme použili neinformativní apriorní rozdělení parametrů – šlo tedy o objektivní bayesovskou analýzu.

²³ Stejně tak intervaly spolehlivosti u frekvenční analýzy obsahují, resp. neobsahují nulu, podle toho, zda je signifikance větší, resp. menší, než 0,05.

6.5.1 Příklad: Lineární regresní analýza

Použitý software BugsXLA nabízí různé typy regresní analýzy. V příkladu provedeme jednoduchou vícenásobnou lineární regresi. Frekvenční analýza je provedena v programu SPSS, bayesovská analýza v aplikaci BugsXLA²⁴. V BugsXLA jsou respondenti, u kterých je alespoň jedna chybějící odpověď, vyřazeni z analýzy (ručně v Excelu), což odpovídá metodě *listwise*, která je součástí frekvenční lineární regrese.

Pro ilustraci jsme vybrali jednoduchý příklad, na kterém budeme zkoumat vliv vzdělání a pohlaví na výši osobního čistého měsíčního příjmu. Závislou proměnnou je tedy spojitá proměnná příjem. Nezávislými proměnnými jsou pohlaví a vzdělání, se kterým nakládáme jako se spojitou proměnnou (má devět kategorií). Cílem je porovnat výsledky regrese pomocí frekvenčního a bayesovského přístupu a ukázat rozdílné výsledky u různých strategií volby apriorního rozdělení. Apriorní rozdělení budeme měnit jen u proměnné vzdělání, tak abychom jednodušeji mohli výsledky porovnat.

Než začneme vytvářet bayesovskou regresi, je třeba specifikovat apriorní rozdělení odhadovaných parametrů. V objektivní bayesovské analýze opět nejdříve ponecháme neinformativní apriorní rozdělení všech parametrů. Poté vytvoříme model zahrnující subjektivní apriorní rozdělení, v tomto případě jde o apriorní rozdělení regresních koeficientů (ne průměrů). Informace o nich jsme získali na základě stejné regresní analýzy provedené na datech z roku 2012²⁵. Pro ilustraci vlivu apriorního rozdělení vytvoříme analýzu za použití mírně subjektivních, značně subjektivních a záměrně chybných apriorních rozdělení.

Apriorní rozdělení proměnných jsme tedy v jednotlivých analýzách nastavili jako:

- a) Objektivní – ponechali jsme defaultní nastavení, tedy pro příjem $\sim N(20000, 1000)$, pohlaví $\sim N(0, 1000)$, vzdělání $\sim N(0, 709, 2)$.

²⁴ SPSS a BugsXLA používají jiné metody odhadu. SPSS v lineární regresi používá metodu nejmenších čtverců, naopak BugsXLA používá metodu maximální věrohodnosti. Výsledky však můžeme bez obav porovnat, neboť „... při předpokladu normality je odhadová funkce nejmenších čtverců pro regresní parametry modelu identická s odhadovou funkcí **metody maximální věrohodnosti (MNV)**.“ (Hušek, 2007, s. 37; zvýraznění v originále)

²⁵ V regresní analýze z roku 2012 mělo vzdělání hodnotu regresního koeficientu 1713 a standardní chybu 166. Ze vztahu mezi standardní chybou a směrodatnou odchylkou (SD je součin SE a odmocniny z N) při počtu respondentů 766 vyplývá směrodatná odchylka 4631.

- b) Mírně subjektivní – rozdělení pro vzdělání volíme jako $\sim N(1700, 31,6)$
- c) Značně subjektivní – rozdělení pro vzdělání volíme jako $\sim N(1700, 10)$
- d) Záměrně chybné – rozdělení pro vzdělání volíme jako $\sim N(0, 10)$

Výsledky všech analýz včetně frekvenční analýzy vidíme v následující tabulce, která srovnává bodové odhady a kredibilní intervaly, resp. intervaly spolehlivosti, regresních koeficientů vysvětlujících proměnných:

Tabulka 14. Výsledky regresních analýz.

| | Bodový odhad | | Intervalový odhad – Pohlaví | | Intervalový odhad – Vzdělání | |
|------------------------------|--------------|----------|-----------------------------|----------------|------------------------------|----------------|
| | Pohlaví | Vzdělání | 2,5 % kvantil | 97,5 % kvantil | 2,5 % kvantil | 97,5 % kvantil |
| a) Objektivní | -5068 | 1713 | -6436 | -3711 | 1383 | 2043 |
| b) Mírně subjektivní | -5067 | 1713 | -6436 | -3711 | 1387 | 2038 |
| c) Značně subjektivní | -5065 | 1704 | -6418 | -3702 | 1534 | 1872 |
| d) Záměrně chybné | -4752 | 433 | -6158 | -3333 | 260 | 607 |
| e) Frekvenční | -5069 | 1713 | -6425 | -3713 | 1388 | 2038 |

Zdroj: CVVM Naše společnost – prosinec 2013. N = 778. Vlastní výpočty.

Na základě tabulky můžeme nyní jednoduše porovnat jednotlivé analýzy. Pro přehlednost se v porovnání zaměříme na proměnnou vzdělání a její regresní koeficient, jehož apriorní rozdělení jsme měnili.

Nejdříve porovnejme výsledky objektivní bayesovské a frekvenční analýzy. Vidíme, že objektivní bayesovská regrese a frekvenční regrese přináší téměř shodné výsledky. Bodové odhady obou regresních koeficientů jsou téměř totožné. Oba intervalové odhady jsou velice široké. Intervalový odhad pro vzdělání je u objektivní bayesovské analýzy o něco málo širší (konkrétně o 10), což je dáno neinformativním apriorním rozdělením.

Objektivní a mírně subjektivní bayesovská regrese přináší co do bodových odhadů shodné výsledky. Díky přesnějšímu (konkrétněji specifikovanému) apriornímu

rozdělení, které jsme proměnné vzdělání přidělili, jsou nyní intervaly spolehlivosti o něco užší. V podstatě jsou co do šířky i hodnot shodné s frekvenčním odhadem.

Dále máme značně subjektivní bayesovskou regresi, ve které je bodový odhad (má hodnotu 1704) více ovlivněn střední hodnotou apriorního rozdělení (ta byla nastavena na 1700) a odlišuje se tak od hodnoty vypočtené ze samotných dat (hodnota 1713). Díky konkrétnější specifikaci apriorního rozdělení pozorujeme, že interval spolehlivosti je v tomto případě podstatně užší.

Nakonec tu máme záměrně chybné apriorní rozdělení, které výsledky odhadů značně vychýlilo. Hodnoty odhadů jsou díky němu zcela nesprávné. Vliv obou proměnných je však stále považován za statisticky významný (vliv by nebyl statisticky významný na hladině významnosti 5 %, pokud by interval spolehlivosti obsahoval nulu).

Můžeme tedy shrnout, že čím přesněji definujeme apriorní rozdělení regresního koeficientu, tím je bodový odhad koeficientu více ovlivněn střední hodnotou apriorního rozdělení a vzdaluje se tak hodnotě zjištěné z napozorovaných dat. Zároveň platí, že čím přesněji definujeme apriorní rozdělení, tím přesnější jsou kredibilní intervaly (zuzují se). Záměrně chybné apriorní rozdělení má i u takto velkého rozsahu výběru značný dopad na výsledky analýzy.

6.6 Shrnutí

Porovnání frekvenčního a bayesovského přístupu můžeme shrnout v následující tabulce:

Tabulka 15. Porovnání frekvenční a bayesovské statistiky.

| Interpretace pravděpodobnosti | |
|--|---|
| F | Frekvenční definice: při velkém počtu opakovaných nezávislých náhodných pokusů lze relativní četnost daného jevu považovat za odhad pravděpodobnosti daného jevu. |
| B | Subjektivní definice: pravděpodobnost jako osobní postoj neboli stupeň důvěry, kterou dává výzkumník nastání daného jevu. |
| Vlastnosti odhadovaného parametru | |
| F | Parametry jsou chápány jako konstanty. Vycházejí z výběrového rozdělení, které vychází ze všech možných výběrů z dané populace. |

| | |
|--------------------------|--|
| B | Parametry jsou chápány jako náhodné proměnné mající pravděpodobnostní rozdělení. Vycházejí z napozorovaných dat (našeho jediného výběru z populace). |
| Odhady | |
| F | Bodové odhady a standardní chyby. Intervaly spolehlivosti, které říkají, že v 19 případech z 20 interval pokrývá skutečnou hodnoty parametru. Pre-data úsudky. |
| B | Popisy posteriorních rozdělení (průměry a kvantily). Kredibilní intervaly, které říkají, že s 95% pravděpodobností se odhadovaný parametr nachází v daném intervalu. Post-data úsudky. |
| Testování hypotéz | |
| F | Na základě předem zvolené hladiny významnosti alfa (obvykle 0,05) a p -hodnoty, což je pravděpodobnost získat již získaný výběrový výsledek či výsledek ještě nepříznivější nulové hypotéze. Pokud je p -hodnota $> 0,05$ nezamítáme nulovou hypotézu. Pokud je p -hodnota $< 0,05$ zamítáme nulovou hypotézu. |
| B | Na základě posteriorního rozdělení a pomocí kredibilních intervalů. P -hodnota znamená pravděpodobnost, že nulová/alternativní hypotéza platí, za předpokladu dat a modelu. |

Na základě porovnání můžeme říci, že z hlediska podstaty obou přístupů je pro sociologické analýzy vhodnější bayesovský přístup. Z něj vycházející analýzy umožňují intuitivní interpretace výsledků, které konvergují se standardním sociologickým porozuměním. Jak už bylo zmíněno, s rostoucím počtem pozorování se výsledky frekvenčního a bayesovského přístupu shodují, nicméně výhody bayesovského přístupu zůstávají zachovány. V sociologii při použití bayesovské statistiky tedy, vzhledem k množství dat, můžeme její výsledky považovat za objektivní. Objektivnost můžeme zajistit také použitím objektivních apriorních rozdělení.

6.7 Software

Bayesovský přístup se používá v mnoha oborech: archeologie, vědy o atmosféře, ekonomie a ekonometrie, vzdělávání, epidemiologie, strojírenství, genetika, hydrologie,

medicína, fyzika, management kvality, sociální vědy (Berger, 2000, s. 1270). Každá z oblastí používá různé metody, pro které bylo tedy vytvořeno mnoho různých programů²⁶, které umožňují analyzovat data bayesovsky.

Jmenujme tedy v této části alespoň čtyři nejdůležitější programy, které umožňují vytvářet mnoho různých analýz a které mohou využít badatelé v sociálních vědách. Základním a nejrozšířenějším softwarem pro provádění bayesovské analýzy dat je BUGS²⁷ (Bayesian inference Using Gibbs Sampling). Z jazyku BUGS jsou dostupné balíčky WinBUGS, který je určen pro operační systém Windows, a OpenBUGS, který je určen pro operační systém Linux. Nevýhodou tohoto programu je to, že pro jeho ovládání je třeba znát jeho poměrně složitý příkazový jazyk. Je však možné k jeho ovládání používat program R²⁸, nebo, jako v případě této práce, lze využít doplněk do Excelu BugsXLA²⁹ (Woodward, 2005 a 2012). Přestože BugsXLA má pro analýzu jen omezené možnosti nastavení i omezené třídy modelů, je přesto pro většinu základních sociologických analýz dostačující. Základní manuál pro tento program uvádíme v příloze této práce. Za zmínku stojí ještě software Mplus³⁰, který je původně určený pro strukturní modelování. Zatímco BUGS, R i BugsXLA jsou zdarma dostupné ke stažení na svých domovských internetových stránkách, program Mplus je třeba zakoupit.

Závěr

Cílem této práce bylo porovnat frekvenční a bayesovskou statistiku v sociologii a porovnat je na analýze konkrétních dat. Tyto cíle byly naplněny, neboť porovnání bylo provedeno jak na teoretické, tak na praktické úrovni. Konkrétně jsme porovnali bodové odhady, intervalové odhady, testování hypotéz a vícenásobnou lineární regresi na reálných sociologických datech.

Díky porovnání analýz můžeme učinit následující závěry. Na základě porovnání bodových a intervalových odhadů jsme zjistili, že výsledky frekvenční a bayesovské analýzy používající neinformativní apriorní rozdělení (objektivní bayesovská analýza)

²⁶ Seznam softwaru vyvinutého po roce 1990 uvádí Berger (2000, s. 1274).

²⁷ <http://www.mrc-bsu.cam.ac.uk/software/bugs/>

²⁸ <http://www.r-project.org/>

²⁹ <http://www.philwoodward.co.uk/bugsxla/>

³⁰ <https://www.statmodel.com/>

jsou téměř shodné. Vliv informativní apriorní informace (subjektivní bayesovská analýza) se vzhledem ke zmenšující velikosti datového souboru zvětšuje, a to tak, že posouvá hodnotu průměru směrem ke střední hodnotě apriorního rozdělení a směrem od průměru spočítaného na základě dat. Také platí, že subjektivní bayesovská analýza poskytuje užší intervalové odhady, než objektivní bayesovská a frekvenční analýza. V případě chybně nastaveného apriorního rozdělení má toto rozdělení jen malý vliv na velký datový soubor (kolem 1000 respondentů), u menších datových souborů (100 a 50 respondentů) jeho vliv roste.

Pomocí porovnání dvouvýběrového t-testu jsme ukázali, že ačkoli bayesovská analýza shrnuje výsledky pomocí posteriorního rozdělení, lze pomocí kredibilních intervalů provést testování hypotéz. V případě frekvenční analýzy se rozhodujeme podle signifikance (a předem stanovené hladiny významnosti), v bayesovské analýze se rozhodujeme podle toho, zda kredibilní intervaly obsahují nebo neobsahují nulu. Na základě toho se pak rozhodujeme, zda testovanou hypotézu přijmeme, nebo zamítneme.

Na příkladu lineární regresní analýzy jsme zkoumali vliv apriorní informace na regresní koeficient jedné z vysvětlujících proměnných. Potvrdilo se, že použití frekvenční a objektivní bayesovské analýzy přináší stejné výsledky. Oproti objektivní bayesovské analýze se u mírně subjektivní a značně subjektivní bayesovské analýzy postupně zužovaly (zpřesňovaly) kredibilní intervaly. Bodový odhad koeficientu se s přesnější apriorní informací mírně vzdaloval od hodnoty vypočtené z dat a přibližoval se střední hodnotě apriorního rozdělení. Přes velký datový soubor (téměř 800 respondentů) mělo záměrně chybné a navíc dosti úzké apriorní rozdělení značný vliv na výsledné odhady.

Můžeme tedy učinit následující závěry:

- a) výsledky frekvenční a objektivní bayesovské analýzy jsou téměř totožné bez ohledu na velikost výběrového souboru,
- b) v případě subjektivní bayesovské analýzy jsou výsledné odhady přesnější – kredibilní intervaly jsou podstatně užší,
- c) pro bayesovské testování hypotéz můžeme využít kredibilní intervaly,
- d) u malých výběrových souborů (cca kolem 50 respondentů) má subjektivní apriorní rozdělení velký vliv na výsledné hodnoty, u velkých výběrových souborů v řádu stovek má subjektivní apriorní rozdělení jen zanedbatelný vliv,

e) chybné apriorní rozdělení má značný vliv na malé datové soubory (do 100 respondentů), pokud je zároveň úzké, má značný vliv i na velké datové soubory (v řádu stovek respondentů).

Výsledky analýz ukazují, že pro sociologii je zvláště vhodná objektivní bayesovská analýza. Sociologové mohou objektivní bayesovskou analýzu téměř bez obav pro své analýzy použít, neboť ani u malých datových souborů nehrozí vliv neinformativní apriorní informace na data. V případě, kdy máme nějakou apriorní představu o hodnotě parametru, je výhodné použít informativní apriorní rozdělení, neboť subjektivní bayesovská analýza přináší přesnější výsledky. Nicméně v případě většiny sociologických šetření vzhledem k velikosti výběru (většinou v řádu stovek či tisíců) je informativní apriorní rozdělení převáženo daty. Přesto je u subjektivní bayesovské analýzy potřeba vybírat apriorní rozdělení pečlivě a ne příliš úzce, neboť chybné a úzké apriorní rozdělení má na výsledky značný a neblahý vliv. V některých případech také hrozí možnost, že přestože budeme vycházet z opakovaných výzkumů používající stejné otázky (proměnné), názor na zkoumaný společenský jev se ve společnosti skutečně změní. Pak by naše apriorní rozdělení, ač přesné, mohlo výsledek ovlivnit negativně. Při použití subjektivní bayesovské analýzy je tedy vhodné vždy provést stejnou analýzu za použití neinformativního apriorního rozdělení a jejich výsledky porovnat a rozhodnout se na základě věcné významnosti.

Kromě výsledných hodnot a jejich přesnosti se na analýzy, které provádíme, můžeme podívat ještě z druhého pohledu, a to filozofického. Na poli teoreticko-filozofickém v případě analýzy sociologických dat jasně vítězí bayesovská statistika. Bayesovská statistika nám díky použití pravděpodobnostního rozdělení pro parametry umožňuje ptát se na to, co nás skutečně zajímá, a interpretovat výsledky intuitivně. S frekvenčními interpretacemi mají obtíže jak studenti, tak výzkumníci. Bayesovská statistika je v tomto ohledu přirozenější, neboť je úzce spjata s výzkumným procesem jako takovým. Začínáme s původní nejistotou o hodnotě parametru. Abychom se o parametru něco dozvěděli, nejprve sesbíráme data. Ve světle nových informací nejistotu o hodnotě parametru zmenšíme. Zatímco frekvenční statistika počítá pravděpodobnost napozorovaných dat (která jsou ve skutečnosti známá a každý se na ně může podívat), v bayesovské analýze se zabýváme nejistotou o hodnotách parametrů, tedy tím, co nás skutečně zajímá. Můžeme souhlasit s Iversenem (1984, s. 76-77), že namísto používání bayesovských interpretací, bychom měli být upřímní a provádět celé analýzy bayesovsky.

Specifikace informativního apriorního rozdělení nám také umožňuje vložit do analýzy znalosti, které o daném jevu máme. V případě subjektivní bayesovské analýzy tak nemusíme začínat vždy „od nuly“ jako frekvenční statistika. To je zvláště vhodné u opakovaných šetření, jako jsou v České republice například kontinuální výzkumy veřejného mínění Centra pro výzkum veřejného mínění Naše společnost nebo mezinárodní sociologické výzkumy jako například *European Social Survey*. Využití bayesovskou statistiku je dále možné také u předvolebních výzkumů volebních preferencí.

Relativní nevýhodou bayesovské statistiky je nutnost znát pravděpodobnostní rozdělení. V případě subjektivní bayesovské analýzy je před vytvořením informativního apriorního rozdělení zapotřebí se nad ním zamyslet. Jak bylo zmíněno na začátku této práce, bayesovská analýza není na rozdíl od frekvenční jako „kuchařka“, ale je potřeba u ní přemýšlet. Pro ty, co jsou ochotni přemýšlet, toto překážkou jistě není.

Největší skutečnou překážkou většího používání bayesovské statistiky je především neexistence jednoduchého programu (jako je ve frekvenční statistice např. SPSS). V tomto můžeme souhlasit s Bergerem (2000, s. 1275), který považuje vytvoření uživatelsky příjemného bayesovského softwaru jako zásadní pro pokrok v používání bayesovských metod. Program BugsXLA použitý v této práci je jistě krokem správným směrem. Jak už bylo naznačeno v kapitole o historii statistiky, někteří významní statistikové předpovídají, že v budoucnosti bude bayesovská statistika získávat stále důležitější postavení. Také podle Hebáka bude vývoj statistiky nejspíš mířit ke kombinování předností klasického a bayesovského přístupu (Hebák, 2012c, s. 86). Není tedy pochyb o tom, že zájem o bayesovskou statistiku bude v budoucnu narůstat. Byla by škoda, kdyby sociologové tento nástup zaspali.

Summary

The goals of this thesis were fulfilled because we compared frequentist and Bayesian analysis both from the theoretical and practical perspective.

In practical perspective we can make these statements:

- a) the outcomes of frequentist and objective Bayesian analysis are identical regardless of the sample size,
- b) by using subjective Bayesian analysis the estimates are more exact,
- c) we can use Bayesian credible intervals for hypothesis testing,

- d) in the case of small sample (around 50 respondents) subjective prior has big impact on outcome, in the case of a big sample (hundreds of respondents) subjective prior has negligible effect,
- e) an incorrect prior has a big effect on small samples (up to 100 respondents), when the prior is also narrow it has a big effect on a big samples (hundreds of respondents) as well.

For sociologists the objective Bayesian analysis is a good way to analyze data because they have numerically the same outcomes as frequentist analysis. In the case we have some prior information about parameter of interest we can use subjective Bayesian analysis so the outcome will be more accurate. Informative prior distribution should be chosen carefully and it is convenient to compare results using non-informative prior distribution.

From the theoretical perspective we can say that for sociological data Bayesian analysis is better than frequentist analysis. Advantage of the Bayesian analysis is that we can find answers for the questions we ask; we can also interpret the results in intuitive way. We can include our prior knowledge into the analysis by specifying informative prior distribution. This can be easily done in repeated surveys and in election surveys.

The disadvantage of Bayesian analysis is that we have to know something about probability distributions and that we have to be willing to think. The only real difficulty that hinders widespread using of Bayesian statistics is the nonexistence of user-friendly software. There is no doubt that in the future Bayesian statistics will play an important role and its impact will grow. Sociologist should keep up with the progress.

Použitá literatura

BAYES, Thomas, Richard PRICE. (1763). An Essay towards solving a Problem in the Doctrine of Chances. By the late Rev. Mr. Bayes, communicated by Mr. Price, in a letter to John Canton, M. A. and F. R. S. *Philosophical Transactions of the Royal Society of London*. roč. 53, s. 370–418.

BELLHOUSE, D. R. The Reverend Thomas Bayes, FRS: A Biography to Celebrate the Tercentenary of His Birth. *Statistical Science*, 2004, roč. 19, č. 1, s. 3-32.

BERGER, James O. Bayesian Analysis: A Look at Today and Thoughts of Tomorrow. *Journal of the American Statistical Association*. 2000, roč. 95, č. 452, s. 1269-1276.

BOLSTAD, William M. *Introduction to Bayesian statistics*. 2nd ed. Hoboken: John Wiley, 2007. 437 s. ISBN 978-0-470-14115-1.

COLEMAN, James S. *Introduction to mathematical sociology*. London: Free Press of Glencoe, 1964. 554 s.

COLLINS, Kathleen M. T. a Anthony J. ONWUEGBUZIE. I Cannot Read My Statistics Textbook: The Relationship between Reading Ability and Statistics Anxiety. *Journal of Negro Education*. 2007, roč. 76, č. 2, s. 118-129.

DAVIS, Jonathan C. a Dan M. SANDIFER-STECH. Wade Into the Water: Preparing Students for Successful Quantitative Research. *Family Relations*. 2006, roč. 55, č. 1, s. 56-66.

DE FINETTI, Bruno. *Theory of probability: a critical introductory treatment. Volume 1*. London: John Wiley & Sons, 1974. 298 s. ISBN 0-471-20141-3.

DE FINETTI, Bruno. *Theory of probability: a critical introductory treatment. Volume 2*. London: John Wiley & Sons, 1975. 375 s. ISBN 0-471-20142-1.

EDLING, Christopher R. Mathematics in Sociology. *Annual Review of Sociology*. 2002, roč. 28, č. 1, s. 197-220.

EDWARDS, A. W. F. [The Reverend Thomas Bayes, FRS: A Biography to Celebrate the Tercentenary of His Birth]: Comment. *Statistical Science*. 2004, roč. 19, č. 1, s. 34-37.

EFRON, Bradley. Bayesian, Frequentists, and Scientists. *Journal of the American Statistical Association*. 2005, roč. 100, č. 469, s. 1-5.

EROLA, Jani. Why Probability Has Not Succeeded in Sociology. *Sociology*. 2010, roč. 44 č. 1, s. 121-138.

FISHER, R. A. On the Mathematical Foundations of Theoretical Statistics. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*. 1922, roč. 222, s. 309-368.

GIGERENZER, Gerd. Mindless statistics. *The Journal of Socio-Economics*. 2004, roč. 33, s. 587-606.

GILL, Jeff. The Insignificance of Null Hypothesis Significance Testing. *Political Research Quarterly*, 1999, roč. 52, č. 3, s. 647-674

GILL, Jeff. *Bayesian methods: a social and behavioral sciences approach*. 2nd ed. Boca Raton: Chapman & Hall/CRC, 2008. 711 s. Statistics in the social and behavioral sciences series. ISBN 978-1-58488-562-7.

HAIGH, John. *Probability: A Very Short Introduction*. Oxford: Oxford University Press, 2012. 144 s. ISBN 978-0199588480.

HALLER, Heiko a Stefan KRAUSS. Misinterpretations of Significance: A Problem Students Share with Their Teachers? *Methods of Psychological Research Online*. 2002, roč. 7, č. 1, s. 1-20.

HEBÁK, Petr a Jana KAHOUNOVÁ. *Počet pravděpodobnosti v příkladech*. 4., nezm. vyd. Praha: Informatorium, 1994. 311 s. ISBN 80-85427-48-6.

HEBÁK, Petr. *Texty k bayesovské statistice*. Vyd. 1. Praha: Vysoká škola ekonomická, 1999. 139 s. ISBN 80-7079-862-9.

HEBÁK, Petr et al. *Vícerozměrné statistické metody*. (1). 2., přeprac. vyd. Praha: Informatorium, 2007. 253 s. ISBN 978-80-7333-056-9.

HEBÁK, Petr. Srovnání klasické a Bayesovské pravděpodobnosti a statistiky (1.). *Acta oeconomica Pragensia: Vědecký sborník Vysoké školy ekonomické v Praze* [online]. Praha: Vysoká škola ekonomická, 2012a, roč. 20, č. 1, s. 69-87 [cit. 2014-04-12]. Dostupné z: <http://www.vse.cz/aop/abstrakt.php3?IDcl=359>

HEBÁK, Petr. Srovnání klasické a Bayesovské pravděpodobnosti a statistiky (2.). *Acta oeconomica Pragensia: Vědecký sborník Vysoké školy ekonomické v Praze* [online]. Praha: Vysoká škola ekonomická, 2012b, roč. 20, č. 2, s. 77-92 [cit. 2014-04-12]. Dostupné z: <http://www.vse.cz/aop/abstrakt.php?IDcl=365>

HEBÁK, Petr. Srovnání klasické a Bayesovské pravděpodobnosti a statistiky (3.). *Acta oeconomica Pragensia: Vědecký sborník Vysoké školy ekonomické v Praze* [online]. Praha: Vysoká škola ekonomická, 2012c, roč. 20, č. 3, s. 76-88 [cit. 2014-04-12]. Dostupné z: <http://www.vse.cz/aop/abstrakt.php?IDcl=371>

HENDL, Jan. *Přehled statistických metod: analýza a metaanalýza dat*. 4., rozš. vyd. Praha: Portál, 2012. 734 s. ISBN 978-80-262-0200-4.

HOFF, Peter D. *A first course in Bayesian statistical methods* [online]. New York: Springer, 2009.

HUŠEK, Roman. *Ekonometrická analýza*. Vyd. 1. Praha: Oeconomica, 2007. 367 s. ISBN 978-80-245-1300-3.

- HUŠKOVÁ, Marie. *Bayesovské metody*. 1. vyd. Praha: Univerzita Karlova, 1985. 93 s.
- IVERSEN, Gudmund R. *Bayesian Statistical Inference*. Sage University Paper Series on Quantitative Applications in the Social Sciences. Beverly Hills, California: Sage Publications, 1984, 80 s. ISBN 08-039-2328-7.
- JACKMAN, Simon. *Bayesian analysis for the social sciences*. Chichester, U.K.: Wiley, 2009. 564 s. ISBN 04-700-1154-8.
- KRUSCHKE, John K. *Doing Bayesian data analysis: a tutorial with R and BUGS*. Burlington: Academic Press, 2011. 653 s. ISBN 978-0-12-381485-2.
- LAZARSELD, Paul Felix. *Mathematical Thinking in the Social Sciences*. Glencoe: The Free Press, 1954. 444 s.
- LECOUTRE, Bruno a Jacques POINTEVINEAU. *The Significance Test Controversy and the Bayesian Alternative* In: Lovric, Miodrag (Editor), (2010). International Encyclopedia of Statistical Science. Heidelberg: Springer Science + Business Media, LLC, reprinted and freely available at StatProb: The Encyclopedia Sponsored by Statistics and Probability Societies.
- LINDLEY, D. V. Fiducial Distributions and Bayes' Theorem. *Journal of the Royal Statistical Society. Series B (Methodological)*. 1958. roč. 20, č. 1, s. 102-107.
- LYNCH, Scott M. *Introduction to applied Bayesian statistics and estimation for social scientists*. New York: Springer, 2007. 357 s. Statistics for social and behavioral sciences. ISBN 978-0-387-71264-2.
- O'HAGAN, A. *Kendall's Advanced theory of Statistics, Vol. 2B, Bayesian Inference*. London: Edward Arnold, 1994.
- PÁZMAN, Andrej. *Bayesovská štatistika*. 1. vyd. Bratislava: Univerzita Komenského, 2003. 98 s. ISBN 80-223-1821-3.
- PRESS, S. James a TANUR, Judith M. *The subjectivity of scientists and the bayesian approach*. New York: John Wiley & Sons, 2001. 274 s. ISBN 0-471-39685-0.
- RAFTERY, Adrian E. Statistics in Sociology, 1950-2000. *Journal of the American Statistical Association*. 2000, roč. 95, č. 450, s. 654-661.
- REITEROVÁ, Eva. Přehled historického vývoje statistiky, její význam v současné době a využití v psychologii. *Acta Universitatis Palackianae Olomucensis. Facultas Philosophica. Varia Psychologica VIII. Psychologica 32*. Olomouc: Univerzita Palackého, 1998. s. 75-86. Dostupné z: <http://publib.upol.cz/~obd/fulltext/psychol8/psychol8-6.pdf>
- ROSENTHAL, Jeffrey S. *Zasažen bleskem: podivuhodný svět pravděpodobností*. Vyd. 1. Praha: Academia, 2008. 294 s. Galileo; sv. 20. ISBN 978-80-200-1645-4.

SOUKUP, Petr. Nesprávná užívání statistické významnosti a jejich možná řešení. *Data a výzkum – SDA Info*. 2010, roč. 4, č. 2, s. 7-104.

WESTERN, Bruce. Bayesian Analysis for Sociologists. *Sociological Methods & Research*. 1999, roč. 28, č. 1, s. 7-33.

WOODWARD, Phill. BugsXLA: Bayes for the Common Man. *Journal of Statistical Software*. 2005, roč. 14, č. 5., s. 1-18.

WOODWARD, Phillip. *Bayesian analysis made simple: an Excel GUI for WinBUGS*. Boca Raton: Taylor & Francis, 2012, s. 342. ISBN 978-1-4398-3954-6.

ŽÁK, Libor. Historie statistiky a pravděpodobnosti. *ÚM FSI v Brně* [online]. 2006, s. 1-4 [cit. 2014-04-12]. Dostupné z: http://mathonline.fme.vutbr.cz/download.aspx?id_file=471

Seznam obrázků

| | |
|--|----|
| Obrázek 1. Proces revize pravděpodobností po získání nových dat..... | 26 |
| Obrázek 2. Doplněk BugsXLA nainstalovaný v Excelu 2003. | 68 |
| Obrázek 3. Specifikace modelu. | 69 |
| Obrázek 4. Nastavení typů proměnných a seřazení hodnot faktorů. | 70 |
| Obrázek 5. Specifikace MCMC algoritmu. | 71 |
| Obrázek 6. Nastavení apriorních rozdělení..... | 72 |
| Obrázek 7. Nabídka výstupů analýzy. | 73 |
| Obrázek 8. Výsledek importovaný do nového sešitu. | 73 |

Seznam grafů

| | |
|--|----|
| Graf 1. Ukázka grafu zobrazujícího apriorní a posteriorní rozdělení parametru. | 43 |
|--|----|

Seznam tabulek

| | |
|---|----|
| Tabulka 1. Odhad průměrného věku pomocí subjektivního apriorního rozdělení v programu BugsXLA (ukázka výstupu)..... | 42 |
| Tabulka 2. Odhad průměrného věku respondenta pro N = 987..... | 43 |
| Tabulka 3. Odhad průměrného věku respondenta pro N = 100..... | 43 |
| Tabulka 4. Odhad průměrného věku respondenta pro N = 53..... | 44 |
| Tabulka 5. Grafy apriorních a posteriorních rozdělení průměrného věku respondenta pro různě velké datové soubory..... | 45 |
| Tabulka 6. Popisné statistiky t-testu v SPSS. | 49 |
| Tabulka 7. Výsledky t-testu v SPSS. | 49 |
| Tabulka 8. Popisné statistiky obdoby t-testu v BugsXLA..... | 50 |
| Tabulka 9. Výsledky obdoby t-testu v BugsXLA..... | 50 |
| Tabulka 10. Popisné statistiky t-testu v SPSS. | 50 |
| Tabulka 11. Výsledky t-testu v SPSS. | 51 |
| Tabulka 12. Popisné statistiky obdoby t-testu v BugsXLA..... | 51 |
| Tabulka 13. Výsledky obdoby t-testu v BugsXLA..... | 51 |
| Tabulka 15. Výsledky regresních analýz. | 54 |
| Tabulka 16. Porovnání frekvenční a bayesovské statistiky. | 55 |

Seznam příloh

Příloha č. 1: Návod k programu BugsXLA (text)

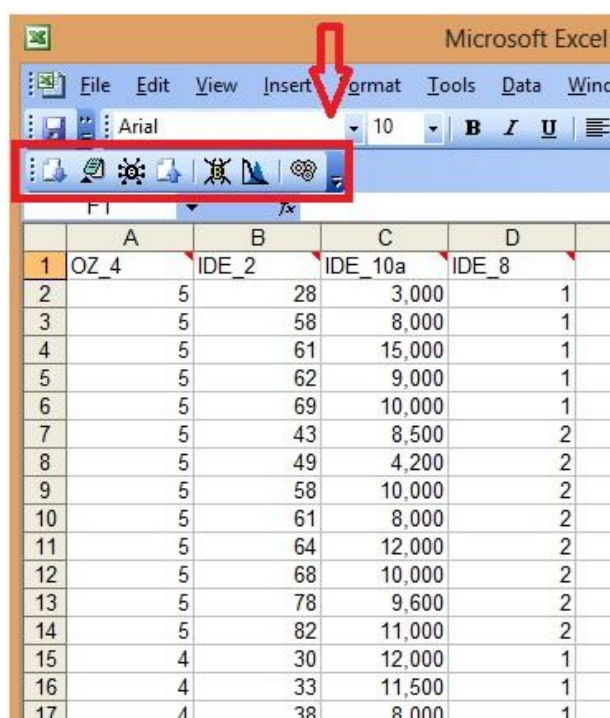
Přílohy

Příloha č. 1: Návod k programu BugsXLA (text)

K tomu, aby byl program správně nainstalován, je potřeba mít nejprve v počítači nainstalován Windows v anglické verzi, Excel v anglické verzi, program Poznámkový blok (resp. Notepad, je součástí Windows) a WinBUGS³¹ (zdarma ke stažení). Následně je potřeba nainstalovat BugsXLA³² (také zdarma ke stažení) a při instalaci postupovat podle přehledného návodu³³ uvedeného na stránkách tvůrce. Poté, co nainstalujeme do Excelu doplněk BugsXLA, můžeme ho začít používat.

Na obrázku níže (Obrázek 2) vidíme, jak vypadají ikony BugsXLA v Excelu (ve verzi 2003). Data pro analýzu je potřeba mít uložená v Excelu tak, aby jednotlivé proměnné tvořily sloupce, jejichž první řádek obsahuje název proměnné.

Obrázek 2. Doplněk BugsXLA nainstalovaný v Excelu 2003.



³¹ <http://www.mrc-bsu.cam.ac.uk/software/bugs/>.

³² <http://www.philwoodward.co.uk/bugsxla/download.html>.

³³ <http://www.philwoodward.co.uk/bugsxla/instr.html>.

Po kliknutí na ikonu *Bayesian Model* (třetí zprava na Obrázku 2) se otevře okno pro zadání specifikace modelu (Obrázek 3). Do kolonky *Data Range* vložíme všechna data (včetně názvů proměnných). Do kolonky *Response*³⁴ se zadává modelovaná proměnná, do kolonek *Factors* a *Covariates* se zadávají nezávislé proměnné.

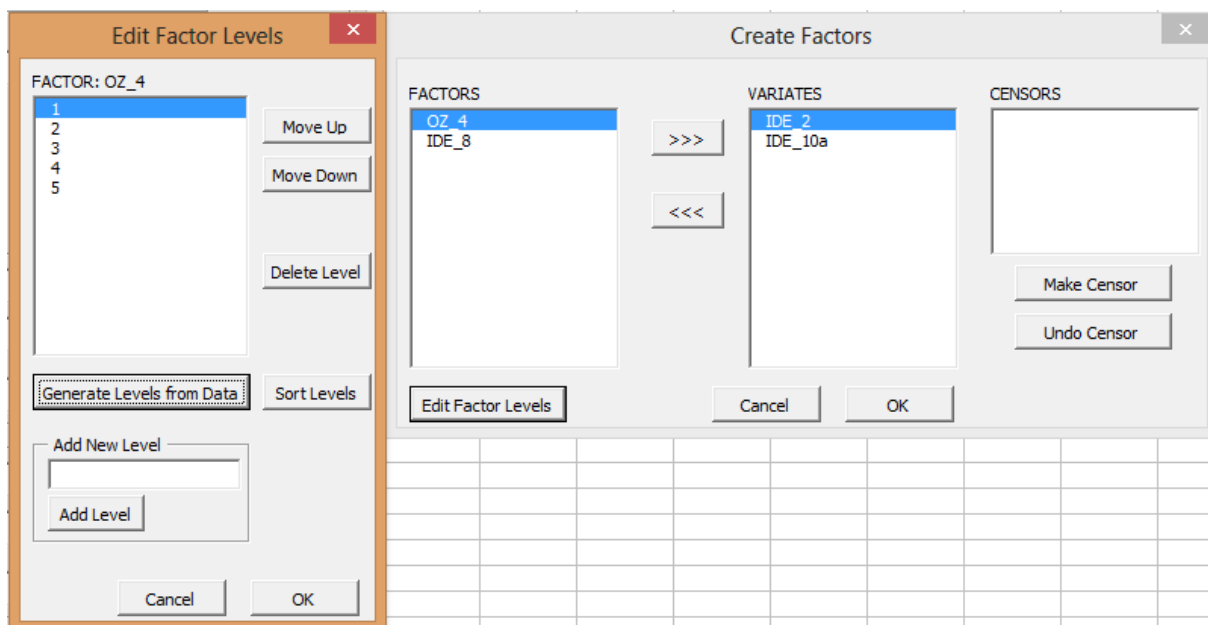
Obrázek 3. Specifikace modelu.

Po vložení dat je potřeba nadefinovat typ proměnných pomocí kliknutím na *Set Variable Types*. Defaultně jsou všechny proměnné považovány za spojité proměnné. V nabídce (Obrázek 4) jako *Factors* nastavíme číselné nebo slovní kategorické

³⁴ Obvykle to musí být náhodná proměnná. Výjimky v zápisu tvoří binomické rozdělení (zapisuje se pomocí lomítka např. Survived/Treated) a Bernoulliho pozorování modelovaná pomocí binomického rozdělení (proměnná musí být nastavená jako *factor* a zapisuje se jako Proměnná/1, kde 1 je první hodnota proměnné)

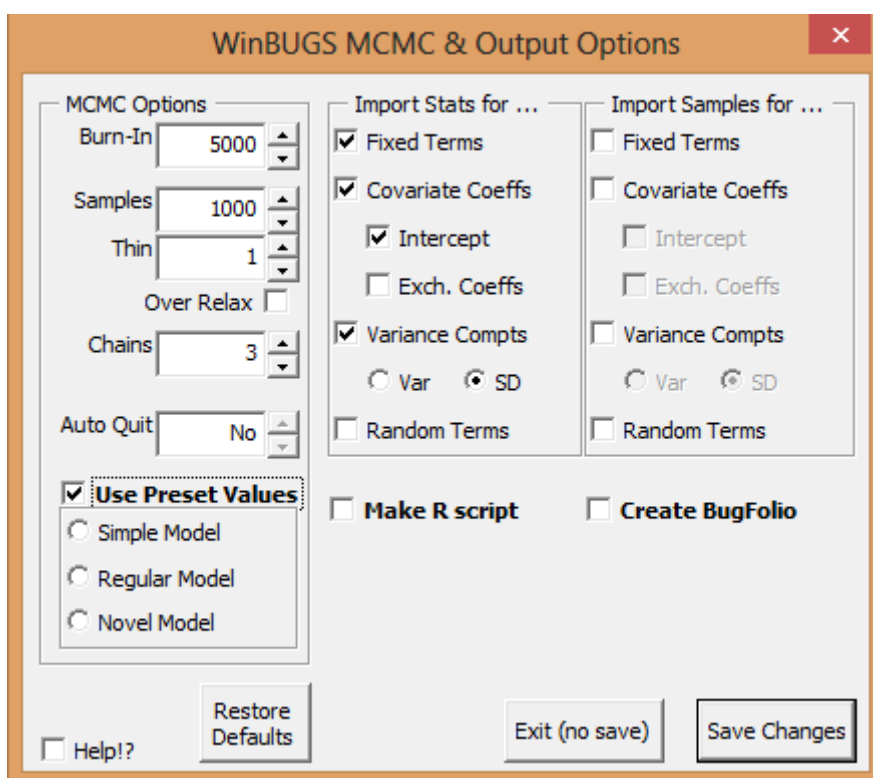
proměnné (nominální nebo ordinální). Jednotlivé hodnoty proměnných definovaných jako *Factors* je nutno seřadit pomocí kliknutí na *Edit Factor Levels*. K seřazení můžeme použít tlačítko *Sort Levels*, nebo je můžeme seřadit ručně (pomocí *Move Up*, *Move Down*).

Obrázek 4. Nastavení typů proměnných a seřazení hodnot faktorů.



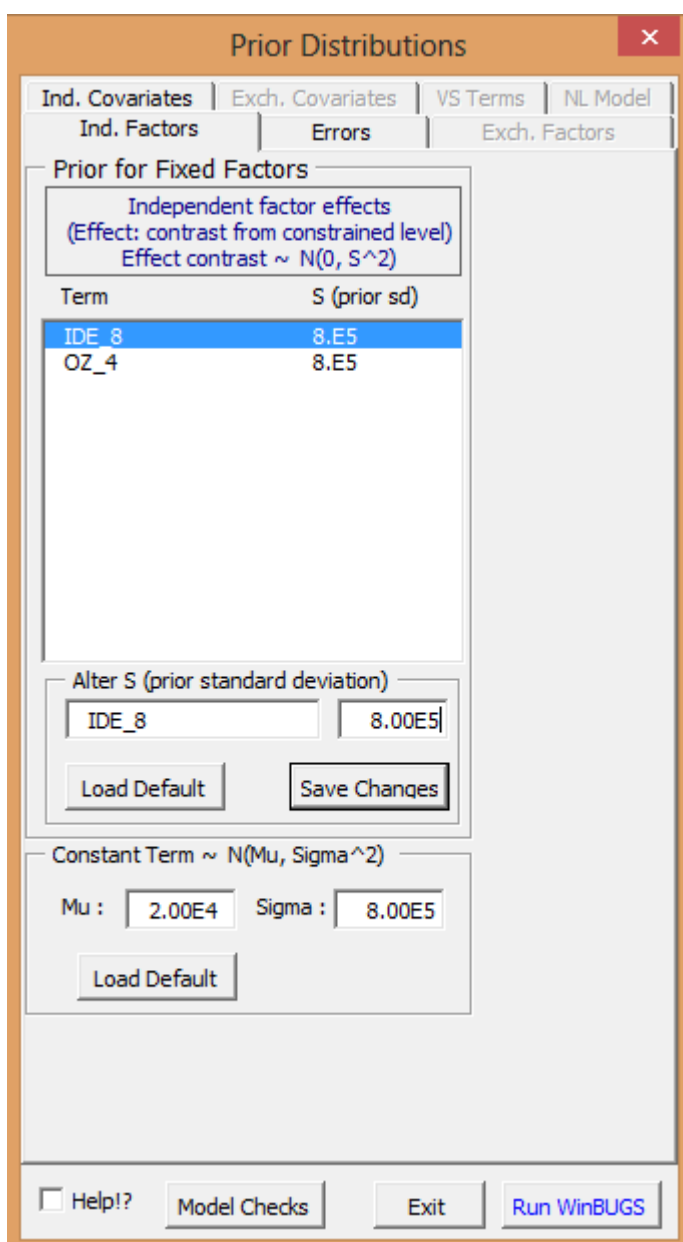
Po specifikování typů proměnných nastavení potvrdíme kliknutím na *OK*, čímž se vrátíme do původního nastavení specifikace modelu (Obrázek 3). Po zadání proměnných do modelu je třeba nastavit MCMC algoritmus pomocí kliknutí na tlačítko *MCMC & Output Options* (na Obrázku 3 vlevo dole).

Obrázek 5. Specifikace MCMC algoritmu.



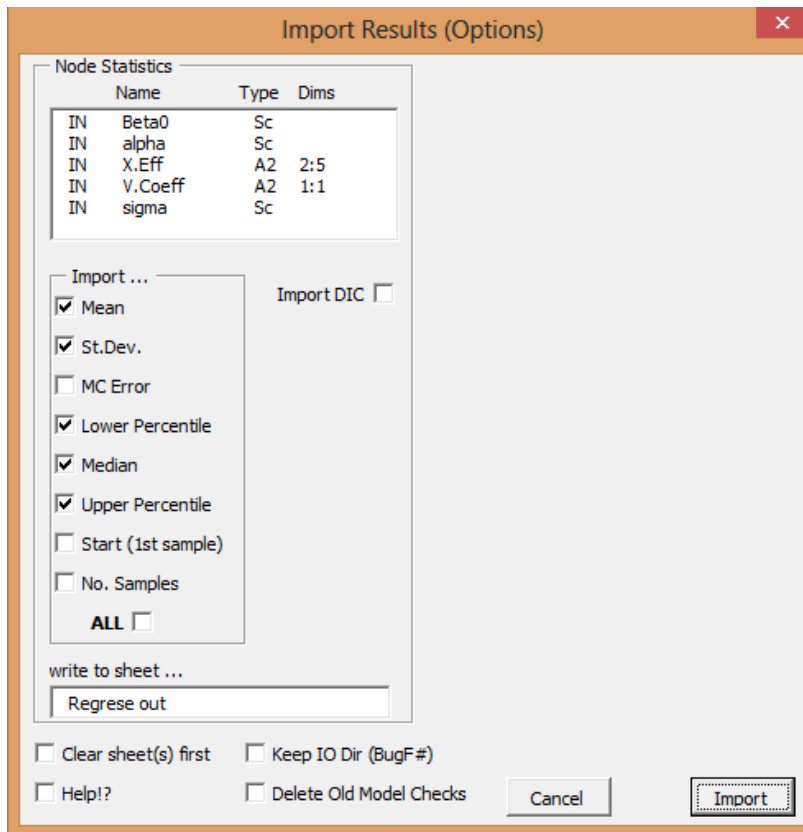
Zobrazí se nám nabídka na nastavení MCMC algoritmů (Obrázek 5). Nastavení *Burn-in* určuje, kolik původních MCMC výběrů má být vyřazeno (po tomto bodě se předpokládá konvergence). *Samples* udává počet MCMC výběrů, které mají být generovány z posteriorního rozdělení. Jeho hodnota se podílí na tom, jak přesně jsou odhadnuty parametry posteriorního rozdělení. Lze využít přednastavených hodnot, které se zobrazí po zaškrtnutí *Use Preset Values*. Zde je důležitá ještě nabídka *Import Samples for* (vpravo nahoře), která nám uloží do sešitu data pro tvorbu grafů srovnávajících apriorní a posteriorní rozdělení parametru. Po nastavení klikneme na *Save Changes*. Tím se vrátíme opět do první nabídky (Obrázek 3). Zde po kliknutí na *OK* vyvoláme nabídku pro nastavení apriorních rozdělení.

Obrázek 6. Nastavení apriorních rozdělání.



Zobrazí se nám karty, na kterých nastavujeme apriorní rozdělání podle typu proměnné (Obrázek 6). BugsXLA nabízí defaultně neinformativní priory (měly by poskytnout shodné výsledky s metodou maximální věrohodnosti) pro všechny parametry. Nastavujeme střední hodnotu a směrodatnou odchylku (u proměnné typu *Factors* je očekávaná hodnota pevně nastavena na nulu). Po nastavení se výpočet spustí pomocí tlačítka *Run WinBUGS*. Tím se spustí WinBUGS, který model spočítá. Výsledné hodnoty se vrátí do Excelu (do nového sešitu – Obrázek 8). Ještě předtím se nás program zeptá, jaké výsledky chceme zjistit (Obrázek 7).

Obrázek 7. Nabídka výstupů analýzy.



Po kliknutí na *Import* se již zobrazí výsledné hodnoty.

Obrázek 8. Výsledek importovaný do nového sešitu.

| | Label | Mean | St.Dev. | 2.5% | Median | 97.5% | WinBUG |
|----|----------------|------------|-----------|------------|------------|------------|------------|
| 2 | CONSTANT | 1.989E+4 | 872.5000 | 1.819E+4 | 1.987E+4 | 2.159E+4 | Beta0 |
| 3 | Intercept at 0 | 1.964E+4 | 1067.0000 | 1.757E+4 | 1.962E+4 | 2.172E+4 | alpha |
| 4 | IDE_8_1 | 0.0000 | 0.0000 | | | | X.Eff[1,1] |
| 5 | IDE_8_2 | -4033.0000 | 537.2000 | -5088.0000 | -4021.0000 | -2968.0000 | X.Eff[1,2] |
| 6 | OZ_4_1 | 0.0000 | 0.0000 | | | | X.Eff[2,1] |
| 7 | OZ_4_2 | -1502.0000 | 911.5000 | -3287.0000 | -1477.0000 | 255.9000 | X.Eff[2,2] |
| 8 | OZ_4_3 | -3610.0000 | 989.9000 | -5514.0000 | -3595.0000 | -1705.0000 | X.Eff[2,3] |
| 9 | OZ_4_4 | -6840.0000 | 1247.0000 | -9275.0000 | -6866.0000 | -4362.0000 | X.Eff[2,4] |
| 10 | OZ_4_5 | -8360.0000 | 2282.0000 | -1.297E+4 | -8355.0000 | -3913.0000 | X.Eff[2,5] |
| 11 | IDE_2 | 5.3430 | 20.6000 | -34.0000 | 4.6590 | 45.3200 | V.Coeff[1] |
| 12 | SD(residual) | 7626.0000 | 193.8000 | 7260.0000 | 7623.0000 | 8015.0000 | sigma |

Note: CONSTANT & Factor effects are determined at the mean of the covariate(s). Interpret these cautiously when Factor x Covariate terms have been fitted.

Model [Regrese[SAS1:\$D\$772]]
 Distribution Normal
 Link Identity
 Response IDE_10a
 Fixed IDE_8+OZ_4
 Covariates IDE_2

Priors
 CONSTANT N(mu=20000, sigma=800000)
 IDE_8 N(mu=0, sigma=800000)
 OZ_4 N(mu=0, sigma=800000)
 IDE_2 N(mu=0, sigma=51200)
 V(residual) Inv-Gamma(0.001, 0.001)

WinBUGS MCMC Settings
 Burn-In: 5000 Samples: 1000 (Thin:1; Chains:3)
 Run took 84 seconds (Auto Quit OFF)
 Run: C:\Program Files\WinBUGS\bin\WinBUGS.exe -c -m -p -r -s -t -u -v -w -x -y -z -aa -bb -cc -dd -ee -ff -gg -hh -ii -jj -kk -ll -mm -nn -oo -pp -qq -rr -ss -tt -uu -vv -ww -xx -yy -zz -AAA -BBB -CCC -DDD -EEE -FFF -GGG -HHH -III -JJJ -KKK -LLL -MMM -NNN -OOO -PPP -QQQ -RRR -SSS -TTT -UUU -VVV -WWW -XXX -YYY -ZZZ -AAA -BBB -CCC -DDD -EEE -FFF -GGG -HHH -III -JJJ -KKK -LLL -MMM -NNN -OOO -PPP -QQQ -RRR -SSS -TTT -UUU -VVV -WWW -XXX -YYY -ZZZ