Ladislav Rabušic, Petr Soukup, Petr Mareš

Statistická analýza sociálněvědních dat (prostřednictvím SPSS)

MASARYKOVA
UNIVERZITA

Ladislav Rabušic
Petr Soukup
Petr Mareš

Statistická analýza sociálněvědních dat (prostřednictvím SPSS)

MUNI
PRESS

Učebnice vychází s laskavou podporou společnosti
IBM Česká republika, spol. s. r. o.

Ladislav Rabušic / Petr Soukup / Petr Mareš

Statistická analýza sociálněvědních dat (prostřednictvím SPSS)

Masarykova univerzita
Brno 2019

KATALOGIZACE V KNIZE – NÁRODNÍ KNIHOVNA ČR

Rabušic, Ladislav, 1954-
Statistická analýza sociálněvědních dat : (prostřednictvím SPSS) / Ladislav Rabušic, Petr Soukup, Petr Mareš. -- 2., přepracované vydání. -- Brno : Masarykova univerzita, 2019. -- 573 stran
Obsahuje bibliografie, bibliografické odkazy a rejstřík

ISBN 978-80-210-9248-8 (vázáno). -- ISBN 978-80-210-9247-1 (brožováno)

* 303.7 * 519.23 * 004.9:311 * 004.42SPSS * 30 * (075.8)
– analýza dat
– statistická analýza
– statistický software
– SPSS (software)
– sociální vědy
– učebnice vysokých škol

311 - Statistika [4]
37.016 - Učební osnovy. Vyučovací předměty. Učebnice [22]

Citace
RABUŠIC, Ladislav, Petr SOUKUP a Petr MAREŠ. *Statistická analýza sociálněvědních dat (prostřednictvím SPSS)*. 2., přeprac. vyd. Brno: Masarykova univerzita, 2019.
ISBN 978-80-210-9247-1 (brožováno)
ISBN 978-80-210-9248-8 (vázáno)
ISBN 978-80-210-9249-5 (online : pdf)

Knihu recenzoval
prof. RNDr. Jan Hendl, CSc.

ISBN 978-80-210-9249-5 (online : pdf)
ISBN 978-80-210-9247-1 (brožováno)
ISBN 978-80-210-9248-8 (vázáno)
ISBN 978-80-210-6362-4 (1. vydání)

Obsah

Kapitola 3

Kapitola 4

Kapitola 5

Kapitola 9

Kapitola 10

Kapitola 11

Kapitola 12

Úvod

Žijeme ve světě, který je prodchnut daty. Data se stala natolik součástí života jedince i společnosti, že se dnes hovoří o datové revoluci. Technologický rozvoj počítačů, rozvoj jejich hardware i software a jejich stále větší schopnost zpracovávat data nejrůznější povahy (data obrazová, jazyková a samozřejmě i numerická) rozšiřuje možnosti vývoje umělé inteligence. Je zřejmé, že schopnost rozumět datům a umět je analyzovat se stává životní nutností, obzvláště u lidí s vysokoškolským diplomem.

Specifickým druhem dat jsou data statistická. Statistická data nás doprovázejí každodenně při čtení novin, poslechu rozhlasu, sledování televizních pořadů. Citování statistických údajů velmi často sloužilo a dodnes slouží v každodenním životě jako důkaz, který má potvrdit správnost argumentace – bohužel v množství nejrůznějších statistických údajů se často nacházejí i takové, které si navzájem protiřečí. To může vést až k jistým pochybnostem o jejich pravdivosti, a potažmo také k pochybnostem o samotné statistice jako vědě (to je o statistické vědě), jak to naznačují dehonestující výroky typu „statistikou lze dokázat cokoliv“ nebo „nevěřím žádné statistice, kterou jsem sám nezfalšoval“. I proto už v roce 1954 napsal americký žurnalista Darell Duff populární útlou knížečku nazvanou *How to Lie with Statistics* (v českém překladu vyšla v roce 2013 pod názvem „Jak lhát se statistikou“) s cílem ukázat, jak se nedopouštět různých druhů chyb (a tedy statisticky nelhat) při intepretaci statistických dat.

Alespoň trochu rozumět statistice, a být tak statisticky gramotný je pro každého nesmírně užitečné. Ostatně již v roce 1950 americký statistik Sam Wilks po zvolení předsedou Americké statistické asociace ve své prezidentské řeči geniálně předpověděl: „Statistické myšlení bude jednou pro efektivní občanství stejně nezbytné jako schopnost číst a psát,“[1] s čímž my, autoři této učebnice, plně souhlasíme; navíc jsme přesvědčeni, že tato doba již nastala.

Pro studenty sociálních věd je základní znalost statistických operací důležitá obzvláště: nejen proto, že jistě chtějí být efektivními občany, ale především proto, že jistě chtějí být i efektivními badateli. A jak již měli možnost v průběhu svého studia

[1] Tento výrok je často mylně připisován známému anglickému spisovateli H. G. Wellsovi, prý pochází z jeho knihy *Mankind in the Making* z roku 1903. Wells se sice v podobném duchu vyjádřil, ale byl to Wilks, jenž Wellsovu myšlenku tak skvěle parafrázoval.

zjistit, značná část sociálněvědních závěrů a generalizujících výroků je založena právě na statistických analýzách. Studenti proto musejí být připraveni na to, že je nutné se statistiku naučit, neboť statistické operace budou organickou součástí jejich výzkumné práce. Proto musejí vědět, že statistické operace jsou založeny na určitých předpokladech, které, pokud nejsou naplněny, vedou k produkci – eufemisticky řečeno – statistických artefaktů, to je – řečeno lapidárně – k produkci mylných výsledků. Ale i ti studenti, kteří se chtějí pohybovat především v prostředí tzv. kvalitativní metodologie, která je založena na „práci bez čísel", by měli považovat zvládnutí základních statistických dovedností za užitečné – přinejmenším proto, aby rozuměli, jak statistické údaje vznikají a jaká čertova kopýtka se ve statistických analýzách mohou vyskytovat.

V této učebnici se budeme zabývat problematikou analýzy statistických dat prostřednictvím softwaru IBM SPSS[2]. Považujeme za důležité hned v úvodu zdůraznit, že čtenářům nepředkládáme klasickou učebnici statistiky (proto také popisujeme základní statistické pojmy, aniž bychom věnovali větší pozornost tomu, jak jsou matematicky definovány), ale soubor návodů, jak statisticky analyzovat datové soubory obsahující hromadné kvantitativní údaje. Učebnice je primárně určena pro studenty společenskovědních oborů. Jako dlouholetí učitelé kurzů „analýza dat" pro studenty sociálních věd totiž máme opakovanou zkušenost, že naučit naše studenty statistické analýze vyžaduje poněkud jiný přístup než prostřednictvím standardní výuky statistiky. Proto je naše učebnice napsána tak, že od čtenáře nevyžaduje více než jen základní znalosti z aritmetiky a elementární algebry. Výklad každé problematiky je řešen podle následujícího vzorce: čtenáři předestřeme analytický problém (například jaká je souvislost mezi mírou religiozity respondentů a jejich postojem k možnosti zavedení eutanazie), poté popíšeme, jakým způsobem se zadá příslušný výpočet v programu SPSS (s názornými návody), a poté ukážeme, jak je možné výsledek, který SPSS vyprodukuje, vyložit a interpretovat. Jelikož všechny analytické úlohy jsou řešeny prostřednictvím výpočtů na počítači, nemusí se nic počítat ručně. Aby ovšem čtenáři pracovali „statisticky poučeně", výkladům některých principů statistiky se samozřejmě nevyhneme. Snažili jsme se ovšem, aby tento výklad byl maximálně srozumitelný, takže jsme museli v mnoha případech výrazně zjednodušovat (a někdy jsme se přitom dostali, jak nás ve svých posudcích upozorňovali naši recenzenti, až na tu nejspodnější možnou hranici zjednodušení). Pevně věříme, že čtenáři pomohou mnohé obrázky, které v knize i výuce hojně užíváme.

Společenské vědy studují, jak známo, sociální jevy (fenomény), to je lidské kolektivní jednání, které je výsledkem vztahů a interakcí mezi lidmi a které se odehrává v prostředí lidské kultury a jejích organizací a institucí. Při jejich poznávání se, jako všechny vědy, řídí třemi cíli: studované jevy se 1) nejdříve musejí **popsat,** poté 2) se musejí prostřednictvím nalezení pravděpodobnostních nebo příčinných (kauzálních) vztahů

[2] Učebnici lze zcela jednoduše užívat též se softwarem PSPP, který je freeware obdobou SPSS (s výjimkou kapitol 14 a 16 a dále některých speciálních postupů v kapitolách 7, 9, 11 a 12). Obdobné nabídky s SPSS i má další freeware nazvaný JASP, jeho analytické možnosti jsou ale ještě omezenější.

vysvětlit a nakonec 3) je třeba se pokoušet o **predikci** (předpověď) budoucího způsobu (popřípadě variantních způsobů) jejich chování nebo existence. Jelikož sociální vědy potřebují k těmto cílům data, povolávají k jejich naplnění ve svém kvantitativním paradigmatu statistickou vědu. Ta umí prostřednictvím svých postupů, to je prostřednictvím postupů **deskriptivní statistiky**, především data **sumarizovat**, tedy **popsat**. Řekneme-li například na základě sociologického výzkumu, který byl proveden na výběrovém souboru 1 812 osob, že v roce 2017 bylo v ČR 90 % respondentů ve svém životě *šťastných*, zatímco *nešťastných* bylo pouhých 10 %,[3] pak jsme statisticky shrnuli 1 812 individuálních odpovědí na otázku, zdali se respondent(ka) cítí celkově šťastný nebo nešťastný. Podobně sumarizující výpovědí bude, když řekneme, že v pocitu štěstí se muži a ženy nelišili nebo že celkově byly v roce 2017 šťastnější spíše mladší věkové skupiny než skupiny starší, neboť ve věku 18–29 let bylo šťastných 95 % respondentů, zatímco ve věkové skupině 60 let a starších bylo šťastných jen 85 %. Postupům této popisné statistiky jsou věnovány především kapitoly 3, 4 a 7.

Hledání vztahů mezi jevy s cílem nalézt jejich pravděpodobnostní nebo kauzální **vysvětlení** jsou věnovány kapitoly 8, 9, 10, 11, 12, 13 a 14. V kapitolách 11–14, jež podávají výklad postupů regresní analýzy, se čtenář navíc seznámí s postupy, které umožňují **predikovat** budoucí chování analyzovaných jevů.

Statistická věda má ovšem pro analýzu sociálněvědních problémů ještě jeden – a to podstatný – moment. Ukazuje, za jakých okolností je možné z údajů, které sociální vědy získají z výběrových souborů (a je pro sociální vědy charakteristické, že ve svých zkoumáních pracují ne s celou populací, ale pouze s její menší či větší částí), zobecňovat prostřednictvím postupů **statistické inference** z dat těchto výběrových souborů na celou populaci (více o tom v kapitole 5).

Jak jsme již uvedli výše, naše učebnice se snaží poskytnout čtenářům pouze základní orientaci v procedurách statistické analýzy. Dobře si uvědomujeme, že tato učebnice z žádného čtenáře statistika neudělá. Věříme ale, že pomůže pochopit, co statistika je, k čemu může sloužit, jak se v ní získávají nejen přesné, ale i spolehlivé a relevantní výsledky, jak těmto výsledkům rozumět a jak je interpretovat – pozor ale, interpretace je již ze značné části za hranicemi znalostí statistiky a musí být vždy doprovázena znalostmi příslušné sociálněvědní disciplíny. Jejím hlavním cílem je tedy naučit, obecně řečeno, jak statistiku používat pro odpovědi na otázky, které si sociální vědy obecně (a sociologie specificky) kladou, a jak přitom udělat co nejméně chybných kroků a rozhodnutí nebo falešných závěrů.

Naše učebnice je výsledkem určité potřeby a z ní plynoucí poptávky, což určuje jak její obsah a výběr jednotlivých témat, tak i rozsah, který jim věnuje. Určuje ale i způsob jejich výkladu. Jsme si přitom vědomi, že se ocitáme v konkurenci se skvělými úvody do statistiky, jako jsou např. *Analýza kategorizovaných dat v sociologii* (Řehák a Řeháková, 1986) nebo *Přehled statistických metod* (Hendl, 2015).

---

[3] Viz Rabušic, Chromková Manea (2018, s. 29) nebo též http://evs.fss.muni.cz/aktualne-k-vyzkumu/vysledky-a-publikace.

Náš výklad je přizpůsoben nejen požadavkům a logice výuky statistiky v bakalářském programu sociologie, ale i logice programu IBM SPSS, který je při ní používán.[4] V české sociologii je právě IBM SPSS momentálně nejužívanější z programů určených pro zpracování hromadných dat. I když, po pravdě řečeno, z možností, které tento program pro statistickou analýzu hromadných dat nabízí, vybírá naše publikace jen malý díl.[5] Z didaktického hlediska upozorňujeme čtenáře na fakt, že pouhá četba našeho textu sice poskytne základní orientaci ve statistice, ale nenaučí jejímu praktickému použití, které je pro svou složitost vázáno na počítačové zpracování hromadných dat. K získání skutečné kompetence při práci s hromadnými daty je třeba při čtení učebnice zároveň pracovat se softwarem a uváděné příklady si krok za krokem skutečně samostatně procvičovat. Z tohoto důvodu může čtenář nalézt všechny datové soubory, s nimiž se v učebnici operuje, na webové adrese: https://iss.fsv.cuni.cz/analyza-dat-spss. Na tento web budou postupně přidávány i další materiály a úlohy pro samostatné procvičování.[6] Ale nejen to, doporučujeme klást si nad příslušnými daty další samostatné výzkumné otázky a zodpovídat si je na základě vlastních výpočtů. Jsme si vědomi toho, že číst učebnici a současně mít otevřený počítač, na němž krok za krokem sledujeme učebnicový výklad tématu a provádíme příslušné počítačové operace, není úplně standardní studijní situace, ale bohužel v tomto případě to jinak nejde. Je to podobné, jako byste se chtěli naučit jezdit na snowboardu. Můžete si o tom, jak se to dělá, přečíst horu návodů, ale dokud to podle nich – a nejlépe s instruktorem – sami nezkusíte, nenaučíte se to nikdy. Ke schopnosti správným způsobem analyzovat statistická (hromadná) data prostřednictvím programu SPSS vede pouze jediná cesta: domácí praktické studium a pravidelné cvičení s učitelem (instruktorem) v rámci výuky. S programem je třeba živě a pravidelně pracovat, platí zde ono okřídlené didaktické *Learning by Doing*, tedy učím se tím, že to sám dělám.

Tato učebnice je novým a rozšířeným vydáním její předchozí verze.[7] S čím konkrétním se čtenář v jednotlivých kapitolách setká? Naše publikace začíná v **první kapitole** odkazy na užitečný metodologický kontext zpracování hromadných dat a pochopitelně se také zabývá otázkami o povaze hromadných dat. Připomínané

[4] Většina výstupů byla generována ve verzi IBM SPSS 22. Víme, že jsou k dispozici již vyšší verze (v době dokončení tohoto rukopisu to byla verze 25), ale podoba výstupů se v jednotlivých verzí nijak zvláště neliší.

[5] Tuto mezeru snad již brzy zaplní připravované vydání pokračování této knihy.

[6] Naším základním datovým souborem, na němž je provedena většina postupů, je reprezentativní soubor České republiky z mezinárodního výzkumu *European Values Study* z roku 1999 (viz soubor EVS99-cvicny). Jsme si vědomi, že tato data jsou pro mnohé čtenáře této učebnice poněkud zastaralá (někteří možná ještě ani nebyli v té době na světě), ale to není z hlediska pochopení smyslu statistické analýzy vůbec důležité. V této knize nám jde primárně o popis a vysvětlení statistických postupů, méně již o věcnou sociologickou analýzu hodnotových orientací a preferencí.

[7] Viz Mareš, P., Rabušic, L., Soukup, P. (2015). *Analýza sociálněvědních dat (nejen v SPSS)*. Masarykova univerzita: Brno. Toto nové vydání nejen rozšiřuje původní texty, ale přidává dvě zcela nové kapitoly o binární a multinomiální logistické regresi. Na druhé straně vypouští kapitolu o analýze dat v prostředí MS Excel. Na webu ke knize budou též postupně publikovány další doplňky a úlohy.

metodologické poznatky jsou sice triviální, ale bez jejich znalosti nelze úspěšně statistiku ve společenskovědním výzkumu používat. Neboť statistika, zdůrazňujeme, je dobrým nástrojem jen pro toho, kdo nejen umí adekvátně aplikovat její postupy, ale současně zná i teorii a metodologii svého vědního oboru. Jen se širšími metodologickými a teoretickými oborovými znalostmi dokážeme formulovat relevantní výzkumné otázky, jsme schopni nalézat relevantní způsoby, jak se pokusit na tyto otázky odpovědět, a nakonec dokážeme vyslovovat relevantní interpretace výsledků našich analýz – relevantních v tom smyslu, že jsou zasazeny do existujícího teoretického kontextu naší disciplíny. Značná část této kapitoly je také věnována připomenutím, jak je pro správnou volbu statistických procedur a konkrétních statistik důležité rozlišovat úroveň měření a typy stupnic použitých proměnných.

Ve **druhé kapitole** seznámíme čtenáře se základními prvky softwaru IBM SPSS. Ukážeme si, jak se se softwarem pracuje, jak se do něj nahrávají výzkumná data, jak vypadá datová matice, jak se její jednotlivé proměnné musejí popsat, aby jim SPSS rozuměl, jak si připravit data pro statistickou analýzu atd. Tato kapitola je poněkud technického rázu, ale bez těchto znalostí není možné se softwarem operovat, a tudíž ani statisticky pracovat.

Třetí kapitola je již věnována prvním krokům statistické analýzy, a to analýze jednorozměrné, která nabízí deskripci rozdělení hodnot jednotlivých proměnných v souboru. Ukazuje, jak zjistit, jaký je podíl jednotek s určitými vlastnostmi ve výběrovém souboru, popřípadě jak jsou v něm jednotlivé vlastnosti rozděleny, což lze vyjadřovat graficky (např. prostřednictvím histogramů) či numericky (prostřednictvím percentilů nebo souhrnných statistik, jako jsou střední hodnoty s jejich mírami variability).

Zvláštní úlohu plní **kapitola čtvrtá**, v níž se věnujeme jednomu ze základních konceptů statistiky, jímž je normální rozdělení. V jejím závěru se při výkladu standardizovaného normálního rozdělení dotkneme problematiky inferenční statistiky a testování hypotéz, kterou detailněji rozvíjí **kapitola pátá**. Ve statistické analýze nám totiž nejde pouze o popis výběrového souboru, s nímž pracujeme, ani o analýzu vztahů mezi proměnnými v tomto souboru. Jde nám o zobecnění výsledků získaných ve výběrovém souboru na populaci, z níž byly jeho jednotky vybrány. Nemohli jsme se proto vyhnout stručnému, a proto snad i poněkud povrchnímu vhledu do počtu pravděpodobnosti. Neboť právě na něm zobecňování (inference) stojí a s ním také padá. Jak konstatoval nositel Nobelovy ceny Ragnar Frische v úvodu ke knize Helmuta Swobody *Moderní statistika*: „… u hypotézy v technickém a statistickém smyslu se soustřeďujeme na charakteristický způsob rozdělení pravděpodobnosti určitého jevu“ (Swoboda, 1971, s. 10). Je důležité si také uvědomit souvislost této kapitoly s kapitolou věnovanou metodologickému kontextu. Inferenční statistika má totiž smysl, jen pokud mají hromadná data určitou povahu. Především musejí představovat takový **výběr** z populace (inference nemá smysl u vyčerpávajících šetření zahrnujících všechny jednotky definované populace), při němž všechny jednotky v populaci mají stejnou šanci, že se do výběru dostanou. A aby naše inference byla korektní, je třeba též vědět, jak určit populaci, z níž náš soubor budeme vybírat. V této kapitole je čtenář také upozorněn na podstatný prvek

práce s výběrovými soubory, totiž že naše výsledky jsou vždy zatíženy tzv. výběrovou chybou a že se pohybují s určitou pravděpodobností v tzv. intervalu spolehlivosti. Tato kapitola se také snaží nabourat mýtus statistické signifikance.

Šestá kapitola je věnována transformacím proměnných, či jinak řečeno, úpravám jejich stupnic. Mnohdy totiž je při sběru dat výhodné použít stupnice, které, pokud nejsou upraveny, mohou analýzu komplikovat, nebo stupnice, které umožní variabilitu úprav podle různě kladených výzkumných otázek. Ať již jde o vhodné slučování hodnot (kategorií) stupnice, změnu orientace pořadových stupnic, aby z nich bylo možno počítat součtové indexy, nebo o různé výpočty s těmito hodnotami. Transformace nám také nejen umožňují úpravy oboru hodnot jednotlivých proměnných, ale umožní nám i vytváření typů kombinacemi hodnot (vlastností zkoumaných jednotek) dvou či více proměnných. Například z hodnot proměnné pohlaví (muž a žena) a dichotomické proměnné pocit štěstí (pocit štěstí a absence pocitu štěstí) lze vytvořit čtyři typy (muž, respektive žena, cítící se šťastnými a muž či žena cítící se nešťastnými).

Ve druhé polovině se učebnice soustředí na základní statistické procedury dvojrozměrné analýzy, která prostřednictvím kontingenčních tabulek hledá vztahy mezi proměnnými a prostřednictvím měr asociace a korelace měří jejich sílu – to je obsahem **kapitol 7, 8 a 9**. A jelikož víme, že v sociální realitě jsou sociální jevy složitě multideterminovány, ukazujeme v **10. kapitole**, jak do analýzy o vztazích mezi dvěma proměnnými přidat působení další, to je třetí proměnné. Na tuto pasáž pak navazují **kapitoly 11, 12, 13 a 14,** v nichž ukazujeme, jak je možné od dvourozměrné deskripce přecházet k vícerozměrné analýze a také k predikci. Tím se dostaneme k tzv. multivariačním (vícerozměrným) analytickým postupům. Poslední dvě kapitoly, **15** a **16**, nás seznámí s tzv. exploračními technikami – faktorovou analýzou (**kapitola 15**) a analýzou seskupovací (**kapitola 16**). K šestnácti kapitolám přidáváme v závěru učebnice **čtyři dodatky**, v nichž naleznou poučení ti, kteří chtějí více a hlouběji zvládnout možnosti, jež nabízí program SPSS.

Na závěr tohoto úvodu si dovolujeme vyjádřit několik poděkování. Náš dík patří především několika generacím našich studentů, na nichž jsme si každoročně postupně ověřovali nové a nové varianty našich textů – dík za to, že to s námi vydrželi a že nám dávali důležité podněty o slabých místech v těchto textech. Dále patří velký dík našemu recenzentu, prof. Janu Hendlovi, za jeho detailní a cenné připomínky k našemu rukopisu – pokud však v textu čtenáři naleznou nedokonalosti, případně i chyby, není to v žádném případě vina recenzentů, ale pouze a toliko vina naše. A budeme pochopitelně našim čtenářům vděčni za jakékoliv podněty pro další zkvalitňování textu.[8]

Ladislav Rabušic, Petr Soukup a Petr Mareš

V Brně a Praze v říjnu 2018

[8] Připomínky prosíme na e-mailové adresy: rabusic@muni.cz a soukup@fsv.cuni.cz.

Kapitola 1

Než začneme

Jasnost je intelektuální hodnota; ne však přesnost a preciznost.
Absolutní preciznost je nedosažitelná; je neúčelné chtít být přesnější,
než to vyžaduje naše problémová situace.

Karl R. Popper

Aforismus o statistice aneb tři druhy lži: lež prostá, lež sprostá, statistika.

Benjamin Disraeli

Memento na začátek

V tomto učebním textu se budeme pohybovat v diskurzu kvantitativního výzkumu, v tzv. kvantitativním paradigmatu. Připomínáme, že sociální vědy jsou vědami multiparadigmatickými, což znamená, že vedle sebe koexistují různé způsoby a pravidla, jak dělat vědu, jak řešit její hlavolamy. Různost těchto vzorců je v podstatě dána tím, jak si jednotlivá paradigmata odpovídají na tři základní otázky: ontologickou, epistemologickou a metodologickou. 1) Ontologická otázka se ptá, jaká je povaha reality, kterou zkoumáme. 2) Epistemologická otázka řeší, jaká je podstata poznání a jaký je vztah mezi tím, kdo poznává, a tím, co je poznáváno. 3) Metodologická otázka se pídí po tom, jakým způsobem se produkuje vědění, porozumění a pochopení. Na tomto základě se dnes definují tři základní skupiny paradigmat: pozitivistické, interpretativní a emancipativní (Mertens, 1998).[9]

Kvantitativní paradigma má svůj vzor v přírodních vědách. Vychází z přesvědčení, že realita je vnější a objektivně poznatelná. Klade velký důraz na měření vlastností, to je na jejich kvantifikaci. Jelikož převážná většina vlastností lidského chování

[9] V literatuře najdeme i další názvy: synonymicky s interpretativním paradigmatem se objevují výrazy etnografické, fenomenologické, hermeneutické nebo naturalistické paradigma. Vedle emancipativního paradigmatu nacházíme také výrazy feministické, participativní nebo kriticky teoretické paradigma.

a lidského světa, jimiž se sociální vědy zabývají, patří ke složitým konstruktům a entitám, musíme se ve výzkumu velmi často spokojit s měřením ne přímo těchto vlastností (nejsou totiž přímo pozorovatelné), ale s měřením jejich pozorovatelných indikátorů. Nemůžeme například přímo změřit vzdělanost jedinců, ale na jejich vzdělanost můžeme usuzovat z výše dosaženého vzdělání. Vzdělanost je v tomto případě vlastností, úroveň dosaženého vzdělání jejím indikátorem.

Nemožnost přímého měření vlastností sociálního světa je v sociálněvědním výzkumu zdrojem jistých potíží. Mnozí metodologové proto zdůrazňují – a my s nimi –, že jedním z klíčových momentů kvantifikace a měření v sociálních vědách je **operacionalizace**, tedy převod abstraktních konstruktů do měřitelných znaků. S operacionalizací je spojen důležitý prvek, a to otázka **validity** těchto operací, což je posouzení, zdali námi vytvořený měřitelný znak (indikátor) je dobrým a skutečným reprezentantem vlastnosti, kterou chceme změřit – proto se také validita definuje jako schopnost měřit to, co skutečně měřit chceme.

Operacionalizace je náročným tvůrčím procesem, v němž postupujeme od sociálněvědních konceptů a jejich nominálních definic přes odhalování jejich dimenzí a subdimenzí ke konkrétním operacím (operacionálním definicím), které nám říkají, co vlastně máme ve výzkumu zjišťovat a měřit. Názornou ukázkou nechť je příklad schématu operacionalizace, které použil de Vaus (1990) pro koncept (pojem) deprivace – viz obrázek 1.1. Co z něj rozhodně stojí za zapamatování, je, že při zjišťování „míry deprivace“ nevystačíme pouze s jedním indikátorem – de Vaus jich navrhuje zjišťovat devět, a to ještě použití škály introverze/extroverze a škály asertivity obsahuje zjišťování řady dalších údajů.[10]

Operacionalizaci a měření vnímáme jako ústřední metodologické téma kvantitativního výzkumu. Jak konstruovat dobré, tedy validní a samozřejmě i dostatečně spolehlivé (reliabilní) měřicí nástroje a jakým způsobem měřit sociální vlastnosti, to jsou kardinální otázky kvantitativního paradigmatu. Podle zastánců paradigmatu kvalitativního jsou ale také zásadními – a z jejich pohledu jen obtížně překonatelnými – překážkami vědecké práce.[11]

Problém měření se navíc umocňuje tím, že velkou část našich kvantitativních dat získáváme na základě standardizovaných výpovědí, to je na základě standardizovaných rozhovorů tazatele se subjekty výzkumu. Standardizovanými výpověďmi rozumíme (i v dalším textu) výpovědi, které lze vyjádřit čísly (věk, příjem apod.) nebo číslicemi (přiřazenými jednotlivým variantám možných odpovědí), umožňujícími

---

4 Zájemce o detailnější vysvětlení problematiky konceptualizace a operacionalizace odkazujeme např. na pasáže v učebnici E. Babbieho (Babbie, 2001, s. 119–145).

5 Není cílem tohoto textu tento spor posuzovat, dodejme pouze, že po letech původně nesmiřitelných diskuzí v poslední čtvrtině minulého století došlo nyní mezi oběma tábory ke smíru a ke koexistenci, především prostřednictvím tzv. *mixed methods research* neboli metod smíšeného výzkumu, který kombinuje (směšuje) relevantní metody kvantitativního a kvalitativního výzkumu. Pro zájemce o metody smíšeného výzkumu může být dobrým úvodem kniha *Advances in Mixed Methods Research: Theories and Applications* editovaná Manfredem Maxem Bergmanem (Sage, 2008).