

Někdy uvažujeme zvlášť proměnnou **pořadovou ordinální**, která vznikne seřazením jedinců podle hodnot získaných změřením nějaké spojité intervalové nebo poměrové proměnné, a **kategoriální ordinální proměnnou**, která rozeznává jenom několik ordinálně uspořádaných kategorií, do kterých jedince nebo objekty zařazujeme.

Poznamenejme, že měřítko znaku spolurozhoduje o tom, jakou statistickou techniku použijeme při zpracování. Často se připomíná odstrašující příklad, jestliže se počítá průměr pro kvantitativně kódovaný kvalitativní znak (např. kódovali jsme žlutý – 0, zelený – 1, modrý – 2). Pak ovšem vede statistické zpracování k nesmyslným tvrzením. Na druhé straně striktní přiřazení statistických technik k určitému typu měřítka vede k nepružnosti analýzy a zbytečnému formalismu. Argumentuje se také, že „měřítko měření“ analyzovaných dat je „empirickou“ otázkou, která je zodpovězena zkoumáním vztahu dané proměnné k ostatním proměnným. Jestliže má vztah lineární formu, údaje je možné považovat za intervalově škálované.

Navíc lze vždy provést transformaci dat s daným měřítkem na data s měřítkem nižší úrovně. Myslí se tím přístup, kdy např. dichotomizujeme metrické spojité údaje pomocí vhodné meze a uvažujeme nový binární údaj, jenž zachycuje dané měření pouze relativně vzhledem k této mezi hodnotami 0 nebo 1 podle toho, zda měření bylo pod, nebo nad zvolenou mezí. Transformaci dat k nižšímu měřítku provádíme často tehdy, jestliže chceme uplatnit neparametrické metody, které vycházejí z četnostních údajů nebo údajů ordinálních. Převod měřítek směrem od nominálního typu k vyšším typům měřítek je méně častý. Vychází se přitom z předpokladu, že většina konstruktů má charakter metrických proměnných, jež jsou navíc v populaci normálně rozložené (viz s. 139).

2.3 Kvalita měření

Při vyhodnocování dat se statistické metody a pravděpodobnostní modely uplatňují ve dvou vzájemně souvisejících oblastech: jednak při analýze dat, která byla získána pomocí specifických měřicích procedur, s cílem zodpovědět v dané oblasti teorie otázky o vztazích mezi konstrukty; ale také při **rozvoji měřicích procedur**. K druhé oblasti teď učiníme několik poznámek.

Proces měření je předpokladem získání dat. Jeho kvalita však není ničím samozřejmým. Před tím, než při zkoumání určitého problému v předmětné oblasti výzkumu analyzujeme data, musíme zaručit, že dokážeme odhadnout vliv kvality měřicích metod na naše výsledky. Kritéria kvality měřicích procedur a metod pro získávání dat se v metodologické literatuře obsáhle diskutují (Blahuš, 1975; Ferjenčík, 2000; Řehák, 1998). Poznamenejme, že jednotlivé oblasti výzkumu

si vytvořily svá pojetí i vlastní terminologii pro zachycení problémů spojených s kvalitou měření. Přírodovědci si např. nebudou ve všem rozumět s psychologem nebo sociologem. Ve všech případech však narážíme na skutečnost, že měřená proměnná odráží jednak vliv vlastního konstruktů nebo měřené veličiny, dále jiných konstruktů (veličin) a náhodné chyby. Jestliže například zjišťujeme hodnoty nějakého psychologického výkonnostního testu, pak mají uvedené komponenty tuto podobu:

- konstrukt, který nás zajímá (např. výkon);
- konstrukty, které nás nezajímají (motivace, strach z testu, jazyková schopnost);
- náhodná chyba (chyby hodnotitele, intraindividuální variabilita, přepisy, chyby přístroje).

Na specifické problémy kvality měření se zaměříme v kapitolách 7.2.9, 7.3.7 a 13.8.3. Zde uvedeme pouze přehled vybraných pojmů. Základní situace je následující: nějakou vlastnost osoby (nebo věci) měříme tehdy, jestliže zkoumané osobě přiřadíme číslo, které reprezentuje tuto vlastnost. Výsledek měření je numerická proměnná, která nabývá různých hodnot v závislosti na tom, jak se lidé nebo věci odlišují. Základní model měření vyjádříme pomocí vztahu mezi neznámou správnou hodnotou proměnné T a naměřenou hodnotou Y

$$Y = T + E,$$

kde E znamená celkovou chybu měření, již lze dále rozložit na několik složek lišících se charakterem i původem. Obvykle rozlišujeme náhodnou chybu a systematickou chybu. Systematická chyba nabývá vždy přibližně stejné hodnoty. Náhodná chyba mění náhodně svoji velikost i směr. Statistickým charakteristikám náhodné chyby říkáme nepřesnost, opakovatelnost nebo reliabilita. Jde nám o to, aby proces měření vedl v rámci daných omezení k dobrému odhadu hodnoty T . Minimalizujeme chybovou část E . Dále zmíníme aspekty, kterými posuzujeme kvalitu měření v behaviorálních vědách. Jestliže máme s danou proměnnou pracovat, snažíme se, aby její měření bylo objektivní, spolehlivé (reliabilní) a validní.

2.3.1 Objektivita

Objektivita měření znamená stupeň toho, jak jsou výsledky nezávislé na výzkumníkovi nebo měřeném jedinci ve smyslu subjektivního úmyslného nebo neúmyslného zkreslení. Při měření fyzikálních veličin v laboratoři se otázka objektivity objevuje zřídka, ale při hodnocení měření v sociologii nebo psychologii se objektivita musí pečlivě přezkušovat.

2.3.2 Spolehlivost (reliabilita)

Spolehlivost (reliabilita) měření znamená stupeň shody (konzistence) výsledků měření jedné osoby nebo jednoho objektu provedeného za stejných podmínek. U testů složených z mnoha položek odpovídá konzistenci hodnot různých podmožin položek mezi sebou.

Nespolehlivost (nízká reliabilita) měření má různý původ. Jeden zdroj nespolehlivosti obvykle nazýváme subjektivní chybou. Zapříčiňuje ji individuální variabilita (únava, klesání zájmu atd.) měřeného subjektu. Pozorovací chyba je jiným zdrojem chyb. Závisí na provedení měření hodnotitelem. Také uvažujeme přístrojové chyby (např. selhání hardwaru).

Existuje mnoho postupů k určení spolehlivosti měření:

- **opakovaná měření** (test-retest reliabilita) – označujeme tak konzistenci neboli shodu opakovaných měření, která jsou oddělena určitým časovým intervalem;
- **měření paralelních testů** – znamená shodu měření s jiným ekvivalentním měřením stejného konstruktů (pokud existují dvě verze A a B téhož testu apod.);
- **půlení testu** (split-half reliabilita) – vyjadřuje, do jaké míry jsou konzistentní jednotlivé části instrumentu měření (nejčastěji se týká různých položek jednoho testu).

V posledním případě se jedná o metodu posuzování internální konzistence, která nevyžaduje u jedince opakované použití měřicí procedury.

Jak se u jednotlivých typů spolehlivosti vypočte numerické hodnocení, stručně uvedeme v kapitole 7.2.9. Pokud měření není spolehlivé, nemůže být ani validní.

2.3.3 Validita

Starší definice **validity** požadovala, aby procedura měření skutečně měřila to, co předpokládáme, že měří. V současnosti se vychází z požadavku, že uživatel má z výsledků měření odvodit správná rozhodnutí. **Validita** odkazuje na přiměřenost, smysluplnost a užitečnost specifických závěrů, jež se provádějí na základě výsledku měření. **Validizace** měřicí metody je procesem k podpoře takového přesvědčení. Posuzují se provedená rozhodnutí, ne měřicí instrument jako takový.

Jak jsme uvedli v předchozím odstavci, platí základní poučka, že bez reliability nemůžeme dosáhnout validity. Metoda měření však může mít velkou reliabilitu, ale přesto malou validitu: „Malá reliabilita zakrývá, nedostatečná validita znetvořuje.“

Koncept validity je triviální, když se jedná o měření znaků jako délka těla. Stává se však komplikovanou záležitostí, jestliže se jedná o osobnostní charakteristiky (např. v psychologii) nebo jiné konstrukty sociálních věd. Při přezkušování validity rozlišujeme obsahovou, kritériální a konstruktovou validitu. Pouze při dostatečném splnění těchto kritérií lze měření rozumně používat a interpretovat.

Při ověřování **obsahové validity** zjišťujeme, do jaké míry měření skutečně reprezentuje dané vlastnosti nebo kvality. Například při konstrukci vědomostních testů si všímáme, zda otázky pokrývají celou problematiku zkoušené látky.

Kritériální validita se posuzuje shodou výsledků zaváděné procedury s nějakou kritériální proměnnou nebo s jiným měřením, které je již ověřené. Ověřenou proceduru měření někdy nazýváme „zlatý standard“. S měřením „zlatého standardu“ srovnáváme výsledky nové měřicí procedury. Uvažujeme *souběžnou validitu*, kdy hodnota kritériální proměnné existuje v současnosti, a *prediktivní validitu*, kdy se hodnota kritériální proměnné realizuje v budoucnu. Například mnoho přijímacích (vstupních) testů užívaných školami nebo personálními agenturami musí být vyhodnoceno ve vztahu k prediktivní validitě.

Konstruktová validita se zabývá teoretickými aspekty měřeného konstruktů (proměnné). Důkazy o konstruktové validitě mohou mít *konvergentní charakter* (test prokazuje vztahy k těm proměnným, jež podle teorie očekáváme) nebo *diskriminační charakter* (naopak nemá vztah k proměnným, když tento vztah neočekáváme). V této etapě přezkušování nové procedury je důležité, zda výsledky predikují stavy, které podle teorie očekáváme. Například test výkonnosti by měl nabývat nízké hodnoty, pokud sportovec nedosahuje dobrých výsledků v soutěžích. Jedna z konstruktových validit se nazývá *zjevná validita* (face validity), jež se zjišťuje na základě prostého úsudku expertů o validitě měření. Konstruktová validita se má vždy ověřovat v rámci daného teoretického kontextu.

■ ■ ■

Pro oblast měření v sociálních vědách bylo navrženo množství měřicích instrumentů. Při používání těchto instrumentů je nutné informovat čtenáře o procesu jejich návrhu a evaluace, o jejich reliabilitě a validitě. Jestliže tato informace schází, pak lze těžko posoudit, zda data mají nějakou informativní hodnotu. Některé studie neinformují o těchto parametrech vůbec, jiné obsahují informaci jenom o reliabilitě nebo jenom o validitě. Protože reliabilita se zjišťuje jednodušeji, autoři se snaží přesvědčit čtenáře, že uvedením reliability demonstrovali i validitu. Oba tyto koncepty mají však odlišnou funkci. Jiné práce zase uvádějí pouze validitu kritériální nebo obsahovou bez zmínky o konstruktové validitě. Předpoklad, že obsahová validita a kritériální validita zajišťují konstruktovou validitu, je také mylný.

Při volbě měřicího instrumentu někdy hrají roli rovněž okolnosti jako cena, obtížnost a časová náročnost realizace, jednoduchost interpretace, dostupnost instrumentu nebo etické aspekty.

Vzhledem k důležitosti celé problematiky kvality měření se tato oblast neustále rozvíjí. Pro hodnocení kvality měřících postupů se využívají komplexní modely a statistické přístupy. Pro souběžné vyhodnocení validity a spolehlivosti měření byly v sociálních vědách navrženy přístupy, jež vyžadují specifické přístupy jak k plánu výzkumu, tak ke statistickému zpracování získaných dat. Například **MTMM** plán (*multitrait-multimethod*) navrhli již v roce 1959 D. T. Campbell a D. W. Fiske.

Na základě měření dvou nebo více konstruktů metoda MTMM hodnotí konvergenční a diskriminační validitu (viz předchozí strana) pomocí analýzy korelací mezi různými měřicími postupy uvažovaných konstruktů. Měření stejného konstruktů by měla silně korelovat, což signalizuje průkaz konvergenční validity. Měření různých konstruktů by měla korelovat slabě, aby byla prokázána diskriminační validita.

Další přístup představuje **teorie zobecnitelnosti** (*generalizability theory*) podle Cronbacha, která využívá především metody analýzy rozptylu a týká se návrhu experimentů a analýzy dat s cílem podrobnějšího prozkoumání jednotlivých chybových složek měření (Cronbach, 1963). Obecná strategie ověřování validity měření zahrnuje i použití kvalitativních (interpretativních) metod. Závěry o validitě nemohou být založeny pouze na statistických metodách. Doporučují se kombinované strategie založené na kvantitativních a kvalitativních metodách.

2.4 Výzkumný plán

Výzkumný plán se zabývá tím, jak má vypadat studie, která umožní zodpovězení výzkumných otázek. V této podkapitole se budeme zabývat především dvěma nejpoužívanějšími typy výzkumných plánů – výběrovým šetřením, které je hlavním zástupcem observačních studií, a experimentem. V **observační studii** sledujeme jedince a měříme proměnné, ale nepokoušíme se ovlivňovat žádnou z proměnných. V **experimentu** cíleně na jedince uplatňujeme intervence nebo ošetření, abychom zjistili, jak ovlivňují závisle (cílovou, odpověďovou) proměnnou.

Při plánování studie kromě validity měření rozlišujeme mezi externí a interní validitou výzkumu. **Externí validita** se týká možnosti zobecnit naše výsledky mimo rámec naší studie. **Interní validita** znamená stupeň průkaznosti studie z hlediska působení nezávisle proměnných na závisle proměnné, jež sledujeme. Týká se toho, do jaké míry můžeme usuzovat o příčinných vztazích mezi oběma typy proměnných. Například při posuzování účinnosti intervence (např. terapie) chceme vědět, zda zlepšení výsledků testů u pokusné skupiny jedinců skutečně způsobila intervence, a ne nějaké jiné příčiny. Externí validita znamená, že podobný efekt se projeví i u jiné skupiny jedinců, u které uplatníme stejnou intervenci. Obecně platí, že větší interní validitu mají dobře provedené experimenty, jež však na rozdíl od výběrových studií založených na náhodném výběru mají

menší externí validitu. Při návrhu výzkumného plánu usilujeme v rámci daných omezujících podmínek o 1. maximalizaci validity měření závisle a nezávisle proměnných; 2. maximalizaci interní validity; 3. maximalizaci externí validity.

2.4.1 Cenzus

Cenzus je metoda sběru dat, v níž se do zkoumání zahrnují všechny jednotky populace. Tento způsob výběru nazýváme také totální výběr. Používá se při sčítání lidu. Totální výběr ale můžeme uplatnit i v jiných souvislostech. Například se pokusíme zjistit názory všech členů akademické obce fakulty.

Sčítání lidu patří mezi nejsložitější statistická šetření kvůli svému rozsahu, ale i k vzhledem k povaze problémů, jež je nutné řešit. Zahrnutí všech bezdomovců nebo členů jiných okrajových skupin představuje obtížný úkol. Cenzus je složitý, časově a finančně náročný projekt. Proto se nahrazuje plány výzkumu, které vycházejí z nějaké formy výběrového šetření.

2.4.2 Výběrové šetření

Ve výběrovém šetření jde o sběr informací standardizovaným způsobem (například pomocí standardizovaného dotazníku) od skupiny lidí. Výzkumník shromáždí data o populaci pomocí nějaké formy výběru jedinců nebo jednotek a sběru hodnot určitých proměnných, ale nepokouší se o jejich ovlivnění. Takové studii říkáme někdy *statistické šetření* nebo *zjišťování* (anglicky *survey*). Sledování skupiny jedinců bez záměrného ovlivňování se říká **observační studie**. Výběrová šetření se považují za podmnožinu observačních studií.

Výběrové šetření je charakterizované dvěma základními vlastnostmi:

- Jedná se o sběr relativně malého množství dat ve standardizované podobě od relativně velké skupiny jedinců.
- Provádí se výběr jedinců z nějaké známé populace.

Tyto vlastnosti se hodí pro popis mnoha šetření, ale existují i případy, kdy shromáždíme o jedinci značné množství údajů nebo ve studii nejde o jedince, ale třeba o organizace jako školy a firmy, tedy dosti rozmanité statistické jednotky. Někdy nemusí být také jasné, jak jsme výběr prováděli – naše šetření se týká dostupné skupiny jedinců. Formální definici výběrového šetření podal Bryman (1989): „Výběrové šetření (survey) znamená shromáždění dat od určitého počtu jednotek, obvykle v jednom časovém okamžiku, s cílem získat systematicky množinu kvantifikovatelných údajů o určitém počtu proměnných, které se pak analyzují, aby se našly vztahy mezi nimi.“ Uvedeme některé důležité způsoby provádění výběru ve statistickém šetření.