

2 Základy statistiky

Se statistikou se setkáváme denně a všude. Naše znalosti o světě vycházejí z velké části ze statistických informací. Statistika je – stručně řečeno – o datech. Data jsou číselné hodnoty, které mají nějaký kontext. Metody statistiky umožňují data znázorňovat a analyzovat vizuálně a numericky, zkoumat jejich struktury, určovat závislosti, dělat o datech závěry. Přitom využíváme naše znalosti kontextu dat.

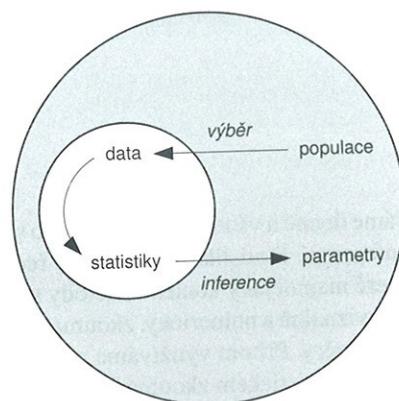
Nejdůležitější otázka ve statistickém zkoumání se týká jistě původu dat. Především rozlišujeme mezi observačními a experimentálními daty. Statistický výzkum začíná přípravou plánu výzkumu, návrhem schématu získávání dat. Pak měříme charakteristiky jedinců nebo objektů, abychom získali čísla, s nimiž budeme pracovat. Ptáme se, zda čísla dávají nějaký smysl. Než je přijmeme, uvažujeme o jejich validitě.

Statistiku lze charakterizovat jako vědu, jež nemá vlastní předmětnou oblast v reálném světě. Je částí teoretické i aplikované matematiky, je také metodologií výzkumu. Jako nauka vychází statistika z abstraktních pojmu, které v konkrétní aplikaci nabývají svůj obsah. S nejdůležitějšími z těchto pojmu se seznámíme v této kapitole.

2.1 Populace, výběr a statistické usuzování

Základní populace je množina všech teoreticky možných objektů (např. jedinců) v uvažované problémové situaci. Například při volbách do parlamentu se zajímáme o všechny osoby s volebním právem v dané zemi. Při zkoumání sportovních jednot v naší republice uvažujeme všechny sportovní jednoty v našem státě. Ne vždy si můžeme populaci takto reálně představit. V mnoha situacích má populace hypotetický význam. Například vytváříme pravděpodobnostní model vývoje nezaměstnanosti. Ten by měl popsat pozorovanou řadu čísel, která je samozřejmě jedinečná. Představujeme si však, že je generována nějakým náhodným mechanismem, jenž dovoluje opakování za stejných podmínek. Za těchto předpokladů lze mluvit o populaci jednotlivých hodnot nezaměstnanosti nebo dokonce o populaci křivek nezaměstnanosti.

Obr. 2.1 Princíp statistického usuzování



Podmnožina základní populace se nazývá **výběr** nebo **vzorek**. Například vybereme určité množství sportovních jednot a podrobíme je zkoumání. Obvykle totiž nemáme možnost z finančních, časových či etických důvodů zkoumat všechny prvky základní populace. Počet prvků n ve výběru nazýváme **rozsah výběru**.

Často chceme pomocí výběru udělat závěr o celé populaci. Proces získávání závěrů o povaze celé populace nebo jejích subpopulací pomocí výběru nazýváme ve statistice **statistické usuzování**. Tento proces je graficky znázorněn na obrázku 2.1.

Postup je následující. Uvažujeme nějakou sledovanou proměnnou u jedinců v populaci. **Populační parametr** proměnné je číselná hodnota, která proměnnou v určitém smyslu charakterizuje. Hledáme např. průměrný roční příjem rodiny v dané oblasti. Populační parametr má nějakou fixní číselnou hodnotu, ale v praxi tuto hodnotu neznáme. Populační parametry se odhadují pomocí **výběrových statistik** neboli prostě statistik. provede se výběr z populace a získá se její vzorek. Vypočítá se **výběrová statistika** pomocí dat získaných ze vzorku populace. Výběrová statistika má číselnou hodnotu, jež popisuje výběr.

Výraz **statistika** tedy užíváme jednak pro matematickou metodu, ale i pro číslo, jež jsme určitým způsobem vypočetli z výběrových dat. Mluvíme tak na jedné straně třeba o *inferenční* nebo *vícerozměrné statistice*, na druhé straně říkáme, že **výběrový průměr je statistika**.

Co je parametr pro populaci, to je výběrová statistika pro výběr. Například chceme odhadnout relativní četnost p specifikované vlastnosti sledovaných jednotek v populaci (neboli počet jedinců s danou vlastností dělený celkovým počtem jedinců v populaci). Taková situace nastává, když sledujeme výskyt jedinců s danou chorobou nebo se snažíme zjistit počet rodin s příjmem pod životním minimem. Pokud zvolíme náhodný výběr sledovaných jednotek, pak relativní četnost vlastnosti ve výběru, označíme ji \hat{p} , je dobrým odhadem čísla p .

U odhadů nás zajímá, jak přesně odhadují výběrové statistiky aktuální parametry celé populace nebo subpopulací. Rozlišujeme přitom:

- chyby výběrové náhodné;
- chyby výběrové systematické;
- chyby nevýběrové.

První dva typy chyb vznikají v důsledku výběru. **Náhodná výběrová chyba** označuje rozdíl mezi výběrovou statistikou a parametrem populace, jenž vznikl v důsledku působení náhody při výběru vzorku populace. Náhodnou chybu odhadu někdy posuzujeme šírkou **intervalu spolehlivosti**, který vytváříme pomocí informací z výběru tak, aby pokryl s velkou pravděpodobností parametr populace. Například jestliže odhadujeme relativní četnost vlastnosti v populaci p pomocí výběrové relativní četnosti \hat{p} , pak tato šířka pro 95% interval spolehlivosti je přibližně $1/\sqrt{n}$, jestliže jsme zvolili výběr náhodně. Jinou metodou je **statistický test hypotéz** o populačních parametrech. Také statistický test pomáhá určit, zda odchylka výběrové statistiky od předpokládané hodnoty parametru populace může vzniknout v důsledku náhodné výběrové chyby. (Podrobněji o těchto metodách pojednáme v kapitole 5 a následujících.)

Systematická chyba poukazuje na zkreslení, jež nemá náhodný charakter. Jde o chybu v plánu výzkumu. Říkáme, že výběr je systematicky zkreslen, pokud při něm dochází k preferování určitých hodnot proměnných. **Chyby nevýběrové** se na rozdíl od náhodné a systematické výběrové chyby mohou vyskytnout nejen při usuzování na základě výběru, ale i při statistickém rozboru dat celé populace (např. aritmetické chyby nebo chyby měřicích instrumentů).

Při identifikaci, kvantifikaci a minimalizaci výběrových chyb s cílem provést **validní statistické usuzování** se začíná návrhem plánu výběru, analýzou jeho vlastností a přípravou. Tyto záležitosti podrobněji probereme v kapitole 2.4.

2.2 Typy proměnných

Statistická analýza se zabývá analýzou dat, která se získávají zjištěním hodnot předem definovaných proměnných, resp. znaků na zkoumaných objektech.

PŘEHLED STATISTICKÝCH METOD

Proměnné nebo **znaky** jsou charakteristiky prvků základního souboru, jež mohou nabývat více hodnot a existuje pro ně předpis, jak tyto hodnoty zjistíme. Například agenturu provádějící předvolební průzkum mohou u voliče zajímat proměnné věk, pohlaví a jakou stranu nebo kandidáta bude volit. Při zkoumání sportovních jednot nás zajímá jejich velikost, počet a typ podporovaných sportů, počet funkcionářů, trenérů a členů, ekonomické ukazatele apod.

Data tvoří aktuální hodnoty proměnných. Některé se zjišťují poměrně snadno – dokážeme např. změřit výšku žáků určité školní třídy. Dokážeme změřit i mnohé jejich psychologické rysy, ale s tím jsou už spojeny určité problémy. Některé proměnné se navíc v čase mění. K určení hodnot mnoha proměnných je zapotřebí značného úsilí. Například zachycení ekonomické situace sportovní jednoty si může vyžádat audit účetnictví, který trvá týden.

Při zkoumání dat mluvíme o rozdelení proměnné. Tímto výrazem rozumíme, jaké hodnoty proměnná nabývá a jak často se jednotlivé hodnoty vyskytují. Pojem rozdelení proměnné pomáhá popsát statistické chování proměnné. Přesnější definici rozdelení proměnné podáme v kapitole 4.

Proměnné lze různě kategorizovat. Popíšeme jejich třídění podle úlohy, jakou hraje daná proměnná ve studii, podle typu použitého měřítka a podle toho, kolik nabývá hodnot.

2.2.1 Závisle a nezávisle proměnné, rušivé proměnné

Většina výzkumných studií rozlišuje **závisle a nezávisle proměnné**. Abychom porozuměli zkoumaným otázkám v dané studii, musíme tyto proměnné správně identifikovat. Někdy označujeme závisle proměnnou *odpověďová, kriteriální nebo cílová*; nezávisle proměnnou nazýváme *prediktor* nebo *explanační proměnná*. Výzkum začíná určením závisle a nezávisle proměnných a další jeho kroky mají určit existenci a charakter jejich vztahu. Často se předpokládá mezi proměnnými příčinný vztah, což znamená, že změna nezávisle proměnné způsobuje změnu závisle proměnné bez ohledu na přítomnost jiných proměnných (podrobnejí Ferjenčík, 2000).

Příkladem závisle a nezávisle proměnných mohou být následující dvojice proměnných, kdy jde o:

- závislost prospěchu (závisle proměnná) na pohlaví žáka v dané věkové kategorii (nezávisle proměnná);
- závislost průměrného prospěchu (závisle proměnná) na počtu žáků ve třídě (nezávisle proměnná);
- závislost síly jedince (závisle proměnná) na objemu jeho svalů (nezávisle proměnná);

- závislost hladiny cholesterolu v krvi (závisle proměnná) na příjmu živočišných tuků (nezávisle proměnná);
- závislost průměrné životní úrovni (závisle proměnná) na podílu zemědělské výroby v celkové ekonomické produkci v dané zemi (nezávisle proměnná).

Při plánování výzkumu a formulování hypotéz je důležité brát v úvahu tzv. rušivé, resp. matoucí proměnné. **Rušivá (matoucí) proměnná** je taková proměnná, která má vztah s cílovou proměnnou a její působení zkresluje naše uvažování o vztahu závisle a nezávisle proměnné. Situaci zjišťování a zohlednění rušivých proměnných nazýváme někdy **problém třetí proměnné**. „Třetí“ rušivých proměnných může být samozřejmě více. V rámci zkoumání mohu být známé nebo neznámé, můžeme je měřit nebo ne, někdy s nimi dokonce můžeme záměrně manipulovat. V některých souvislostech nazýváme rušivou proměnnou „moderující“ proměnná.

PŘÍKLAD 2.1

Problém třetí proměnné

Uvažujme modelovou situaci. Sledujeme závislost rizika úmrtí řidiče při autonehodě na použití ochranných pásků. Posuzujeme četnosti jednotlivých kombinací, jež jsou popsány v tabulce 2.1. Rozdíl mezi riziky úmrtí při použití a nepoužití ochranných pásků je 0,2 (tento rozdíl je také statistický významný, když použijeme příslušný statistický test, viz s. 301). Na první pohled údaje studie demonstrují výhody při používání ochranných pásků. Hlavní problém této tabulky spočívá v tom, že neuvažuje vážnost nehody měřenou např. rychlosí při nárazu.

Tato proměnná byla v modelové situaci zaznamenána. Příslušná data popisuje tabulka 2.2. Tato tabulka vypráví o něčem trochu jiném. Při malých rychlostech je riziko smrti stejné u obou skupin řidičů, totiž 0,1. Totéž platí pro riziko při velkých rychlostech, které je ale mnohem větší, totiž 0,6. Data tedy naznačují, že ochranné pásky nemají žádný vliv na riziko smrti. Skeptický čtenář se může zeptat, zda lze data vysvětlit nějakým jiným způsobem.

Tab. 2.1 Příklad výběru sledovaných proměnných – hypotetická data o autonehodách

	Ochranný pás		Součet
	použitý	nepoužitý	
Řidič zemřel	10	20	30
Řidič nezemřel	40	30	70
Součet	50	50	100
Riziko	0,2	0,4	

Tab. 2.2 Účinek rušivého faktoru – podrobnější data o nehodách

	Nízká rychlosť			Vysoká rychlosť		
	Ochranný pás		Součet	Ochranný pás		Součet
	použitý	nepoužitý		použitý	nepoužitý	
Řidič zemrel	4	2	6	6	18	24
Řidič nezemrel	36	18	54	4	12	16
Součet	40	20	60	10	30	40
Riziko	0,1	0,1		0,6	0,6	

Patrně jde o to, že řidiči při nehodě v malých rychlostech používají ochranný pás mnohem častěji než při velkých rychlostech. To ukazuje relativní četnosti 40/60 a 10/40. Rychlosť nárazu je v tomto případě rušivým faktorem, protože maskuje efekt – pokud je nějaký – faktoru „použití ochranného pásu“. Vliv rušivého faktoru vede ke zkreslení odhadu efektu hlavní nezávisle proměnné.

Pro skutečná data první pohled naznačuje, že pásy snižují riziko smrti o 60 %. Po podrobnější analýze beroucí v úvahu další proměnné, např. fakt, že lidé, kteří řídí nepřipoutání, jezdí i v jiných ohledech nebezpečněji a mají vážnější nehody, různé výzkumy dospěly k závěru, že pásy snižují pravděpodobnost úmrtí „jen“ o 40–50 % vzhledem k nepřipoutané osobě.



„Jistě, Vaši pacienti mají o 50 % méně zubních kazů. Ale to je proto, že mají o 50 % méně zubů!“

V daném výzkumném kontextu známe a sledujeme jenom některé rušivé proměnné – ty někdy nazýváme kovarianční proměnné. Existují tři hlavní metody, jak se vypořádat s působením rušivých proměnných:

- Při jejich znalosti je můžeme zohlednit při návrhu plánu výzkumu (např. metodou vytváření bloků nebo vyrovnáváním čili mačováním – podrobněji kap. 2.4.3).
- Omezení působení neznámých potencionálně matoucích proměnných docilujeme výběrem homogenních skupin pomocí mechanismu randomizace. Tím docílíme, že hodnoty jsou ve všech skupinách stejně zastoupeny a jejich vliv na cílovou proměnnou stejný ve všech skupinách (viz kap. 2.4.3).
- Známé rušivé proměnné můžeme změřit a jejich působení eliminovat při statistickém zpracování (podrobněji kap. 7.2.4, 8.5 a 10.7).

■ ■ ■

U nezávisle proměnných rozlišujeme mezi manipulativním a přirozeným (organismickým) prediktorem. **Přirozený prediktor** nemůžeme ovlivňovat, je atributem jedince, objektu nebo prostředí. Jedná se např. o inteligenci měřenou prostřednictvím IQ, výšku, délku slunečního svitu, pohlaví. **Manipulativní prediktor** můžeme měnit.

V některých projektech je také nutné zohlednit různé typy proměnných podle hierarchie statistických jednotek, jichž se týkají. Rozlišujeme tak individuální proměnné a kolektivní proměnné. Mezi **individuální proměnné** patří věk jedince, jeho vzdělání apod. **Kolektivní proměnné** charakterizují podskupinu, do níž jedinec patří (např. typ výuky ve třídě). Další úroveň hierarchie určuje typ školy, obce apod. Tato typologie proměnných se využívá v tzv. víceúrovňové analýze, jejíž podstatu objasňujeme v kapitole 13.4.

2.2.2 Proměnné podle typu použitého měřítka

Pojem měření se často ztotožňuje – pod vlivem přírodních věd – se získáváním metrických údajů. Naměřená hodnota je v nich násobkem nebo částí jednotky měření. V sociálních vědách se používá definice, která říká, že **měření** znamená přiřazení čísel objektům, jež je vázáno na teorii daného oboru. Operace přiřazování se řídí určitými pravidly.

Proměnné lze rozlišovat podle toho, co jejich hodnoty vypovídají. Mluvíme pak o **škále měření** nebo použitém **měřítku**, jež charakterizuje, nakolik je zobrazení vlastnosti do množiny reálných čísel (způsob, jakým přiřazujeme různým hodnotám proměnné čísla) plnohodnotné vzhledem k operacím mezi čísly.

Kvalitativní (nominální) měřítko znamená přiřazení, které pouze vyjadřuje, že lze rozlišit jednotlivé hodnoty. Například žáky lze rozlišit podle toho, zda umějí plavat, nebo ne. U těchto měřítek můžeme určit počet použitých kategorií. Jestliže rozlišuje použité měřítko jenom dvě třídy, mluvíme o binárním (alternativním, dichotomickém) znaku, jinak o polytomickém znaku.

Ordinální měřítko kromě rozlišení tříd ještě vyjadřuje nějaké jejich řazení podle intenzity nebo pořadí. Data s ordinálním měřítkem lze uspořádat. Například můžeme seřadit žáky podle jejich prospěchu z matematiky.

Intervalové měřítko má vlastnosti ordinálního měřítka. Navíc vyhovuje požadavku, že vzdálenosti jednotlivých údajů uvnitř tohoto systému jsou dány nějakou jednotkou měření. Proto je možné je sčítat a odčítat.

Měřítko z největší informační hodnotou je **poměrové měřítko**, které má všechny vlastnosti intervalového měřítka a navíc vlastnost existence absolutního nulového bodu.

Obě posledně uváděná měřítka společně označujeme jako metrická, a spolu s ordinálním měřítkem je shrnujeme do skupiny **intenzivních měřítek**.

Vymezíme popsaná měřítka pomocí matematických symbolů. Na dvou objektech A a B získáme měření s hodnotami x_a a x_b . Pro jednotlivá měřítka platí následující pravidla:

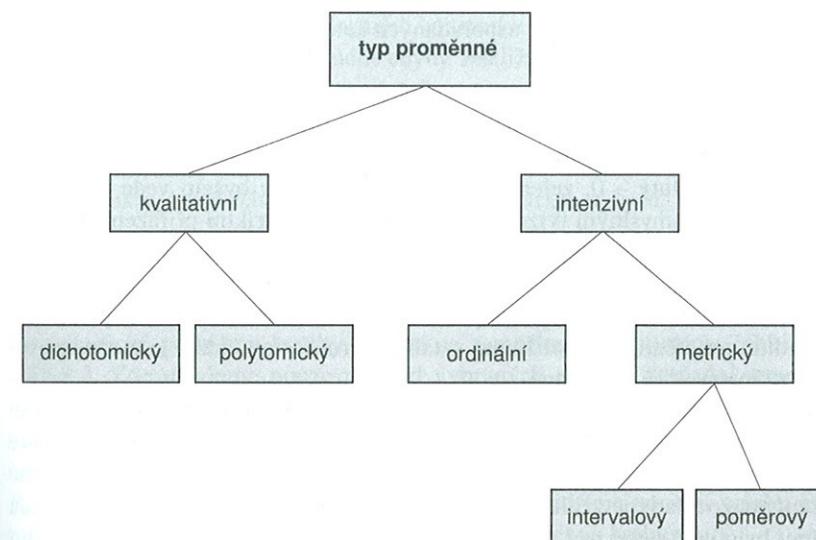
- Při *nominálním měřítku* můžeme mít pouze vztahy $x_a = x_b$ nebo $x_a \neq x_b$ (např. muž, žena).
- *Ordinální měřítko* také dovoluje vztahy $x_a > x_b$ nebo $x_a < x_b$ (např. světlý, šedý, tmavý).
- *Intervalové měřítko* navíc předpokládá, že může být definována velikost rozdílu, takže objekt A se liší o $x_A - x_B$ jednotek od objektu B (např. teplota ve stupních Celsius).
- *Poměrové měřítko* má navíc definovanou absolutní nulu, takže má smysl říci, že A je x_a/x_b větší než B, jestliže $x_a > x_b$ a $x_b \neq 0$ (např. teplota v Kelvinově stupnici).

Ve statistice nehráje rozdíl mezi intervalovým a poměrovým měřítkem velkou roli. Statistické metody však zohledňují rozdíl mezi nominálními, ordinálními a metrickými typy údajů. Klasifikaci proměnných podle typu měřítka znázorníme graficky (viz obr. 2.2). V tabulce 2.3 uvádíme některé příklady různých typů proměnných.

2.2.3 Diskrétní a spojité proměnné

Při zpracování dat hraje rovněž to, zda jsou diskrétní, nebo spojitá. **Spojitá proměnná** může teoreticky nabývat libovolných hodnot z určitého intervalu reálných čísel. **Diskrétní proměnné** neboli kategoriální proměnné nabývají naopak

Obr. 2.2 Přehled typů proměnných podle použitého měřítka



Tab. 2.3 Příklady různých typů proměnných

Proměnná	Typ
pohlaví	kvalitativní dichotomický
provozovaný sport	kvalitativní polytomický
prospěch	ordinální kategoriální
teporá frekvence	metrický spojitý i diskrétní
počet vstřelených gólů	metrický diskrétní
rodinný stav	kvalitativní polytomický
věk	metrický spojitý

pouze konečného počtu hodnot (pro matematiky můžeme dodat: nebo nejvýše spočetně mnoha hodnot). Podle kontextu zpracování lze i s diskrétními údaji pracovat jako se spojitými, např. tepovou frekvenci můžeme považovat za spojitou proměnnou, pokud se setkáváme v analýze s velkými rozdíly tepových frekvencí.

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 rozehnává jenom několik ordinálně uspořádaných kategorií, do kterých jedince nebo objekty zařazujeme.

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 kvalitatívní 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 povážovat za intervalově škálované.

žovat 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 měřítka 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.

procedur. K druhé oblasti tehdy učiníme nekonk poznání. Proces měření je předpokladem získání dat. Jeho kvalita však není nížním 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. Přírodnovědci si např. nebudou ve všem rozumět s psychology nebo sociology. Ve všech případech však narázíme na skutečnost, že měřená proměnná odráží jednak vliv vlastního konstruktu 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 problémky kvality měření se zaměříme v kapitolách 7.2.9, 7.3.7 a 13.8.3. Zde uvedeme pouze přehled vybraných pojmu. 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 je objektivita musí pečlivě prezkušovat.

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 podmnožin položek mezi sebou.

Nespolehlivost (nízká reliabilita) měření má různý původ. Jeden zdroj nespolehlivosti obvykle nazýváme subjektivní chybou. Zapříčinuje ji individuální variabilita (únavu, 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:

- **opakování měření** (test-retest reliabilita) – označujeme tak konzistenci nebo shodu opakových měření, která jsou oddělena určitým časovým intervalem;
- **měření paralelních testů** – znamená shodu měření s jiným ekvivalentním měřením stejného konstruktu (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 opakové 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 přiměřenosť, 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 reliabilitu 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, kriteriální 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 všimáme, zda otázky pokrývají celou problematiku zkoušené látky.

Kriteriální validita se posuzuje shodou výsledků zaváděné procedury s nějakou kriteriální 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 kriteriální proměnné existuje v současnosti, a *prediktivní validitu*, kdy se hodnota kriteriální 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 konstruktu (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 neoceníme). 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 chází, 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 reliabilita demonstrovali i validitu. Oba tyto koncepty mají však odlišnou funkci. Jiné práce zase uvádějí pouze validitu kriteriální nebo obsahovou bez zmínky o konstruktové validitě. Předpoklad, že obsahová validita a kriteriální validita zajišťují konstruktovou validitu, je také mylný.

PŘEHLED STATISTICKÝCH METOD

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 měřicích 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 kriminační validitu (viz předchozí strana) pomocí analýzy korelací mezi různými měřicími postupy uvažovaných konstruktů. Měření stejného konstruktu by měla silně korelovat, což signifikuje průkaz konvergenční validity. Měření různých konstruktů by měla korelovat slabě, aby byla prokázána diskriminační validita.

Další přístup představuje **teorie zobecitelnosti** (*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 (cifrovou, odpověďovou) proměnnou.

Při plánování studie kromě validity měření rozlisujeme 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 příčinách 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 dobré 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 shromažďuje 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 shromažďujeme o jedinci značné množství údajů nebo ve studii nejde o jedince, ale řeč 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, obyčejně v jednom časovém okamžiku, s cílem získat systematicky množinu kvantifikovatelných údajů o určitém počtu proměnných, které se pak analyzují, aby se nalezly vztahy mezi nimi.“ Uvedeme některé důležité způsoby provádění výběru ve statistickém šetření.

Způsoby výběru

Jsme v situaci, že není možné (třeba z finančních důvodů) získat data od celé populace. Chceme proto provést výběrové šetření. Plán výběru označuje metodu, která se použije pro výběr podmnožiny jedinců nebo objektů ze základní populace. Popíšeme čtyři typy výběrových plánů.

- **Výběr na základě dobrovolnosti** se často aplikuje v průzkumech veřejného mínění. V tomto plánu se obvykle jedná o získání odpovědí na jednu nebo několik otázek. Jedinci z populace se sami rozhodují, zda odpovědět, nebo ne. Například po televizním pořadu jsou diváci vyzváni, aby se vyjádřili k diskutovaným otázkám. Pravděpodobně na výzvu reagují pouze vysoce motivovaní diváci.
- **Výběr na základě dostupnosti.** Jedinci jsou z populace vybráni na základě dostupnosti a výhodnosti. Například jestliže provádíme průzkum o nákupních zvykách jedinců ve sportovní prodejně, zvolíme výběr 100 jedinců na základě okamžité dostupnosti v této prodejně v danou sobotu dopoledne. Jiný příklad představuje medicínský výzkum s pacienty, kteří se léčí v dané nemocnici.
- **Kvótní výběr.** V této proceduře výběru mají tazatelé za úkol provést rozhovor s určitým počtem jedinců v několika různých kategoriích obyvatelstva. Za kategorie se volí např. věk, pohlaví nebo ekonomický status. Vychází se obvykle z demografických informací o obyvatelstvu.
- **Náhodný výběr.** Základní doporučení říká, že je pro statistické šetření nejlepší. Představuje však ideál, který je v praxi často jen nedostatečně uskutečněn, resp. neuskutečnitelný. Náhodný nebo pravděpodobnostní výběr ze základní populace je takové vytvoření její podmnožiny, kdy:

 1. každý prvek populace má známou pravděpodobnost, že bude do výběru zařazen;
 2. výběr je proveden pomocí metody, jež tyto pravděpodobnosti výběru realizuje;
 3. pravděpodobnosti výběru prvků se uvažují při zpracování získaných dat.

PŘÍKLAD 2.2

Metody výběru při výběrovém šetření

Výbor sportovního klubu, který má 1000 členů, se rozhodne dotázat 50 členů na spokojenosť se službami. K jakému typu přiřadíme následující metody výběru?

■ Dotáže se prvních 50 členů, kteří přijdu daný den ráno do klubu.

■ Nechá na pultu ve vrátnici dotazníky a počká, až členové při vstupu vyplní 50 dotazníků.

■ Jména všech členů se dají do tomboly a vylosuje se 50 členů, kteří pak vyplní dotazník.

První tři popsané techniky výběru – výběr na základě dobrovolnosti, dostupnosti a kvótní výběr – nejsou zcela optimální, protože získaná data mohou být zkreslena. Zkreslení je systematická chyba, která se projeví, jestliže data se získávají od jedinců, kteří patří do subpopulace se specifickými znaky. Výběr na základě dobrovolnosti preferuje jedince, kteří jsou nějakým způsobem motivováni k tomu, aby se k dotazům vyjádřili. Dobrovolníci mají obvykle více extrémní názory. Výběr na základě dostupnosti zahrnuje skupinu dostupnou na daném místě a v daném okamžiku. Je možné, že zákazníci v sobotu dopoledne se liší v nákupních zvykách od ostatních zákazníků. Hlavním nedostatkem kvótního výběru je subjektivní volba jedinců z dané kategorie jedinců. Není jisté, zda tazatel nevědomky nedává přednost určitému typu respondentů. Prostý náhodný výběr eliminuje výběrové zkreslení, protože všechny podmnožiny daného rozsahu mají stejnou šanci, že budou vybrány pro pozorování, dotazování nebo měření.

Prostý náhodný výběr

Základní typ pravděpodobnostního výběru je **prostý náhodný výběr** – pravděpodobnostní výběr, kdy každý prvek populace má stejnou pravděpodobnost, že se do výběru dostane. Prostý náhodný výběr je také definován jako výběr o rozsahu n , kdy každá množina n prvků má stejnou pravděpodobnost, že bude vybrána.

Abychom takový výběr realizovali, musíme mít k dispozici očíslovaný seznam všech prvků základního souboru – tzv. **oporu výběru**, a dále generátor náhodných čísel, pomocí něhož vybereme očíslovaný prvek z opory výběru. Předpokládáme, že populace má N prvků a výběr má mít n prvků. Procedura výběru se sestává z následujících kroků:

1. sestavíme seznam jednotek celé populace (opora výběru) a přiřadíme jako identifikaci každé jednotce celé číslo od 1 do N ;
2. rozhodneme se, jak bude veliký rozsah výběru n ;
3. vygenerujeme n náhodných celých čísel mezi číslem 1 a N ;
4. získáme data od jedinců identifikovaných v opoře výběru těmito náhodnými čísly.

Poměr mezi rozsahem výběru (n) a velikostí populace (N) nazýváme výběrový poměr:

$$\text{výběrový poměr} = \frac{\text{rozsah výběru } n}{\text{velikost populace } N}$$

Například vybíráme-li náhodně 3 jedince z populace o velikosti $N = 600$, pak výběrový poměr je $3/600 = 0,005$ nebo $0,5\%$. Tento poměr vyjadřuje pravděpodobnost, že prvek základního souboru je zařazen do výběru. Výběr se může dělat s vrácením nebo bez vrácení. Jestliže prvek vrátíme do základního souboru,

má nenulovou pravděpodobnost, že bude do výběru vybrán vícekrát. Výhodnější pro statistické odvozování různých formulí je výběr s vracením. V takovém případě je však vhodné, aby výběrový poměr byl malý (< 5 %).

V mnoha případech nevíme, zda byl náš výběr proveden jako náhodný. Pak předpokládáme, že pracujeme s množinou prvků získanou prostým náhodným výběrem a doufáme, že náš předpoklad není příliš nesprávný. S takovou situací se nejčastěji setkáváme, když vytváříme výběr na základě dostupnosti – zkoumají se např. všichni pacienti, kteří se objevili v ordinaci v daném časovém intervalu. Zde nastává obtíž s vymezením, pro jakou základní populaci vlastně naše výsledky budou platit. Důkladná diskuse charakteristik takové množiny pacientů dovolí posoudit, zda je možné výsledky použít i v jiných kontextech.

Náhrada prostého náhodného výběru

Stává se, že prostý náhodný výběr je neproveditelný nebo nákladný, hlavně tehdy, když je základní populace značně rozsáhlá. Uvádíme některé přijatelné náhradní metody výběru, jež využívají ve výběru náhodný mechanismus.

- **Stratifikovaný náhodný výběr.** Jestliže populace obsahuje různorodé subpopulace, je možné rozdělit populaci do těchto skupin a provést prostý náhodný výběr pro každou skupinu. Tyto podskupiny se nazývají strata nebo vrstvy. Podskupiny jsou voleny tak, že jsou více homogenní. Výsledky pro všechny skupiny pak tvoří výběr. Tato technika je vhodná, jestliže populaci lze stratifikovat podle pohlaví, věku nebo demografických parametrů a výzkumník chce zajistit reprezentaci každé podskupiny.
- **Vícestupňový shlukový výběr** se často používá pro získání informací o věrném mírnění. Například chceme zjistit názory lidí z panelových sídlišť měst určité velikosti. Postupuje se např. takto: 1. vybere se náhodně vzorek okresů; 2. z takto vybraných okresů se v každém okresu vybere náhodně určitý počet měst o dané velikosti; 3. pro takto vybraná města se vybere náhodně vzorek z jejich sídlišť; 4. z vybraných sídlišť se náhodně vyberou domácnosti, ve kterých se provede dotazování. V každé vrstvě shluků se provádí náhodný výběr. Tato vícestupňová procedura vypadá velmi komplikovaně, ale ve skutečnosti je velmi efektivní a méně nákladná než prostý náhodný výběr domácností ze sídlišť. Pro analýzu získaných dat se používají speciální techniky. Někdy však lze aplikovat postupy určené jinak pro zpracování dat z prostých náhodných výběru.
- **Systematický výběr.** Tato metoda začíná soupisem a očíslováním prvků populace. Pak se provede rozhodnutí, jak z tohoto seznamu systematicky vybírat prvky. Například se bude vybírat vždy jeden prvek z padesáti. Nejdříve vybírat prvky. Například se bude vybírat vždy jeden prvek z padesáti. Nejdříve

se zvolí náhodně prvek z první padesátky. Ten má určité pořadí. V dalším kroku se k tomuto číslu postupně přičítává číslo 50 a prvky s takto získaným pořadovým číslem se zařazují do výběru. Aby byl systematický výběr validní, musí výzkumník zajistit, že číslování prvků není závislé ne charakteristikách jedinců, jež se mají zkoumat.

PRÍKLAD 2.3

Systematický výběr při výběrovém šetření

Vraťme se k příkladu o sportovním klubu. Vedení chce provést výběr o rozsahu 50 z celkem 1000 členů, aby zjistilo spokojenosť členů klubu. Zvolí následující strategii: Jednotlivé členy v seznamu označí čísla od 1 do 20 tak, že prvky seznamu postupně očíslouje touto sérií čísel jejím opakováním použitím. Náhodně vybere celé číslo z intervalu 1 až 20. Pak se dotáže všech členů s tímto označením. Jedná se o systematický výběr, který je založen na pravděpodobnosti, ale prostřednictvím jiného mechanismu, než je tomu u prostého výběru. Při tomto schématu nemá každý jedinec stejnou pravděpodobnost, že se do výběru dostane. Nemůžeme např. získat výběr, jenž bude současně obsahovat jedince s označením 10 a 11.

Typy výběrových šetření

Uvedeme definice často používaných plánů výběrového šetření.

- **Jednorázové průřezové šetření** je hlavní metodou sociologického výzkumu. V jednom časovém bodu je podroben zkoumání výběr cílové populace. Abychom demonstrovali složitost výběrového šetření, uvádíme jednotlivé fáze exploračního jednorázového výběrového šetření spolu s odhadem časových nároků pro studii s rozsahem výběru 1000 dotazovaných jedinců.
 1. Počáteční příprava. Definice účelu a určení specifických požadavků na informace. Návrh typu výběru a způsobu sběru dat – sběr dat lze provést pomocí sítě tazatelů, poštou, telefonicky nebo elektronicky (6 týdnů).
 2. Konstrukce dotazníku a manuálu pro dotazování, včetně pilotní studie (minimálně 20 jedinců) (6 týdnů).
 3. Podrobné interview jedinců v pilotní studii pro přezkoušení odpovědí (6 týdnů).
 4. Provedení dotazování (6 týdnů).
 5. Kódování, oprava dat a převedení dat do počítače (6 týdnů).
 6. Analýza a interpretace dat, příprava zprávy (2 měsíce).

Mentální šetření vyžaduje menší zdroje a lze ho realizovat za kratší dobu. Probíhá však ve stejných fázích, přičemž navrhování dotazníku zabere relativně nejvíce času.

PŘEHLED STATISTICKÝCH METOD

K dalším typům výběrových šetření patří:

- **Panelové šetření** je opakované šetření, prováděné na stále stejném souboru osob vybraných z cílové populace. Tento soubor se nazývá panel.
- **Opakované průlezové šetření** je opakované šetření, které se podobá panelovému tím, že opakovaně zjištujeme údaje z cílové skupiny. Pokaždé však jde o nově vybraný soubor jedinců z cílové skupiny.
- **Kontrolovaná prospektivní studie** je observační výzkum, v němž se delší dobu sleduje více skupin jedinců vybraných výzkumníkem, aby se zjistil výskyt určité události (např. nemocnost, úmrtnost). Na začátku studie a v jejím průběhu se zjišťují hodnoty všech důležitých explanačních proměnných. Příkladem jsou jedinci rozděleni do dvou skupin podle toho, zda jsou exponováni (vystaveni) posuzovanému rizikovému faktoru (s expozicí – kuřáci, bez expozice – nekuřáci). Sledujeme několik let vývoj jejich zdravotního stavu a přítomnost definovaných událostí (výskyt zhoubného nádoru).
- **Kontrolovaná retrospektivní studie** je observační studie, v níž se nejdříve zjišťuje hodnota cílové proměnné a pak se zkoumá historie jedinců, aby se odhalily relevantní vysvětlující proměnné. Přitom pracujeme také s paralelní historickou skupinou, u níž nebyla zjištěna specifikovaná hodnota cílové proměnné. Této studii se také říká studie případů a kontrol (např. případy jsou jedinci se zhoubným nádorem, kontroly jsou jedinci bez zhoubného nádoru).

Oba posledně jmenované plány pracují delší dobu s kohortami. **Kohortu** definujeme jako skupinu lidí, která nabyla podobnou zkušenosť ve stejném čase. Někdy tyto studie nazýváme kohortové longitudinální studie.

Výhody a nevýhody výběrového šetření

Výzkumníci mívají dosti polarizované názory na výhody a nevýhody výběrových šetření. Někteří je považují za ústřední v každém empirickém sociálním výzkumu. Jiní oponují, že pomocí dotazníkových šetření a jím podobných akcí výzkumu generuje ohromné množství dat pochybné hodnoty a že takové studie právě stírají falešnou exaktnost kvantitativním hodnocením jevů; poznatky se většinou opírají o zkreslující informace od jedinců, jejichž odpovědi jsou směsi zdvořilosti, nudy, neznalosti nebo předstírání. Některé výhody a nevýhody shrnuje tabulka 2.4.

Příčiny zkreslení výběrového šetření

Existuje mnoho potenciálních příčin, které vedou ke zkreslení výsledků výběrového šetření. Při hodnocení kvality studie zkoumáme různé aspekty studie:

Tab. 2.4 Výhody a nevýhody výběrového šetření

	Výhody	Nevýhody
Obecně	Výběrová šetření pomáhají relativně jednoduše a přímočaře při studiu postojů, názorů a motivů členů populace ve standardizované podobě.	Data jsou ovlivněna vlastnostmi respondenta (pamětí, znalostmi, motivací). Respondenti nechtějí bezpodmínečně informovat o svých skutečných názorech a pocitech.
Poštovní dotazování a jedincem vyplňovaný dotazník	Často se jedná o jedinou snadnou cestu, jak získat informaci od většího počtu jedinců rychlou cestou. Dovoluje anonymní vyplňování, které podporuje upřímnost, jestliže se jedná o citlivé informace.	Obvykle slabá návratnost vyplňených dotazníků. Chybné odpovědi nelze opravit. Respondent nemusí odpovídat pravdivě.
Dotazování pomocí tazatelů	Tazatel může vyjasnit otázky. Přítomnost tazatele zvyšuje angažovanost účastníka, tazatel může posoudit, zda účastník odpovídá seriozně.	Data mohou být ovlivněna vlastnostmi tazatele a interakcí mezi respondentem a tazatelem. Respondent může mít pocit, že jeho odpovědi nebudou anonymní.

- **Kdo prováděl šetření?** Výzkum preferencí voličů může provádět např. jedna politická strana. Neovlivní to získané výsledky?
- **Jaká byla populace?** Koho jsme se ptali?
- **Jaká byla opora výběru?** Pokud opora výběru nezahrnuje celou populaci, výběr bude zkreslen stejným způsobem, jako je zkreslená opora výběru. Například pokud použijeme jako oporu výběru očíslované jedince z telefonního seznamu, jde o zkreslenou oporu výběru, jež nemusí adekvátně reprezentovat celou populaci. Efekt „nepokrytí“ nastane, když některé skupiny nebyly z nějakého důvodu zahrnuty do výběru.
- **Jak byl vzorek vybrán?** Byl použit náhodný výběr?
- **Jak byl vzorek veliký?** Je důležité vědět rozsah výběru a dosažené odhadu chyb v podobě intervalů spolehlivosti.
- **Jaká byla návratnost?** Kolik dotazovaných subjektů vrátilo vyplňený dotazník?
- **Jak byli jedinci kontaktováni?** Telefonicky, poštou nebo osobním rozhovorem doma?
- **Kdy bylo šetření provedeno?** Stalo se tak po nějaké významné události, která může ovlivnit odpovědi?
- **Došlo ke zkreslení odpovědí?** Je dán špatným sestavením dotazníku nebo subjektem dotazovaným, jenž úmyslně dává falešné odpovědi. Dotazovaný také může být ovlivněn negativně tazatelem.
- **Jaké otázky se přesně kladly?** Závisí na tazateli.

Výběrové šetření a příčinnost

Důležitý rozdíl mezi výběrovým šetřením a experimentem spočívá v hodnotě získaných závěrů o příčinách (kauzálních) vztazích. Pokud je experiment správně navržen a proveden, můžeme usuzovat na *příčinné působení* mezi ovlivňující a cílovou proměnnou. Správně provedená výběrová studie poskytuje pouze informaci o závislosti mezi ovlivňující (explanační) proměnnou a odpověď (cílová proměnná). Doklady o kauzálním vztahu jsou v tomto typu výzkumného plánu velmi slabé.

Ve výběrových studiích musíme pátrat po možném působení rušivých proměnných. Například ve výběrové studii o kouření získáme mnoho informací o možném negativním působení kouření. Avšak zástupci tabákového průmyslu budou tvrdit, že pozorované efekty jsou způsobeny nesledovanými rušivými proměnnými, jako jsou genetické predispozice, působení prostředí, stravovací zvyky atd. Také věku je možné přisoudit vlastnost rušivé proměnné. Kontrola ne-působení rušivých proměnných pomocí experimentu je však v tomto případě nemožná, protože takový pokus by byl neetický. Lze však provést výběrové šetření, kdy vybereme pokud možno dvě stejné skupiny kuřáků a nekuřáků.

Na příkladu otázky důkazu příčinného vztahu mezi kouřením a rakovinou plic demonstруjeme základních **pět kritérií příčinnosti**, jejichž platnost musíme přezkoumávat, abychom mohli vyslovit věrohodný názor na podstatu sledovaného vztahu.

1. **Asociace mezi proměnnými je silná.** Asociace mezi kouřením a úmrtností na rakovinu plic je silná. Sílu pozorované asociace lze měřit některým z koeficientů korelace.
2. **Asociace má být konzistentní.** Mnoho studií různých skupin lidí v mnoha zemích prokázalo spojení mezi kouřením a rakovinou plic. Tato skutečnost znamená, že rušivá proměnná působící v jedné skupině vysvětlí nelenzenou asociaci.
3. **Větší dávka má větší účinek.** Lidé, kteří kouří více cigaret a delší dobu, mají častěji rakovinu. Ti, kdo s kouřením přestanou, toto riziko zmenšují.
4. **Předpokládaná příčina předchází účinek.** Rakovina plic se vyvíjí delší dobu. Počet lidí s rakovinou plic se zvyšuje s nárůstem obliby kouření.
5. **Příčina je teoreticky vysvětlitelná.** Existují propracované teorie, které modelují vznik rakoviny jako důsledek kouření.

Shrnutí důležitých aspektů výběrového šetření

Výběrové šetření má za cíl statistický popis trendů, postojů nebo názorů v populaci pomocí zkoumání výběru z populace. Z výběru provádí výzkumník zobecnění nebo úsudky o celé populaci. V tomto typu výzkumu nemáme žádný vliv na sledované proměnné. Návrh provedení výběrového šetření se drží zavedeného formátu. V odborných časopisech nalezneme mnoho příkladů použití této strategie výzkumu. Uvádíme kontrolní otázky, které vyjasňujeme jak při navrhování výběrového šetření, tak při kritickém posuzování prezentovaných výsledků:

1. Je určen účel výběrového šetření?
2. Jsou uvedeny důvody pro volbu určitého plánu šetření?
3. Jaká je povaha šetření (průřezové, longitudinální)?
4. Je vymezena populace a její velikost?
5. Bude se populace stratifikovat?
6. Kolik lidí se bude zkoumat, jak bude veliký vzorek? Jak se došlo tomuto číslu?
7. Jaká se zvolí procedura pro realizaci výběru (náhodná, nenáhodná)?
8. Jaké se použijí instrumenty pro získání údajů? Kdo je navrhl?
9. Jaká obsahová oblast se bude zachycovat? Jaké se použijí škály?
10. Jak se ověří navržený postup? Provede se pilotní studie?
11. Jaký je časový rozvrh celé akce?
12. Jaké jsou proměnné ve studii? Odpovídají všechny proměnné účelu studie?
13. Jsou určeny kroky analýzy, aby bylo možné:
 - analyzovat počet zadaných a vyplněných dotazníků,
 - kontrolovat systematickou chybu v odpovědích,
 - provést popisnou analýzu,
 - vypočítat hodnoty škál,
 - kontrolovat spolehlivost vyplnění odpovědí,
 - provést případné inferenční kroky při statistické analýze?

2.4.3 Experiment

Experiment je studie, v níž výzkumník pomocí záměrných změn podmínek zkoumá, jaké změny nastaly u jedné nebo více skupin jedinců nebo jiných homogenných jednotek. Změna podmínek spočívá např. v provedení určité intervence nebo ošetření. Tzv. přirozené experimenty využívají nezáměrných změn podmínek. Jestliže sledujeme rozdílnosti dvou nebo více skupin, mluvíme o komparativním experimentu. Pro popis změn podmínek, které má výzkumník pod kontrolou, využíváme pojem nezávisle proměnných. Výzkumník sleduje, jak

PŘEHLED STATISTICKÝCH METOD

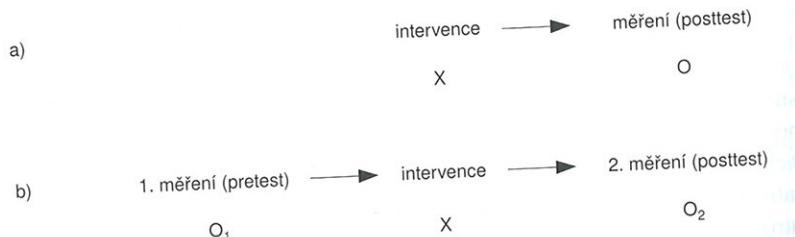
různé hodnoty nezávisle proměnných, jejichž hodnoty záměrně volí, ovlivňují cílovou sledovanou proměnnou (závisle proměnnou). Statistickými metodami posuzujeme, zda odchylky mezi skupinami vznikly v důsledku náhody, nebo ne.

Experiment se provádí na sledovaných jednotkách. Lidského jedince v této souvislosti nazýváme **subjektem**. Je důležité zdůraznit, že experimentátor zařazuje jednotky do různých skupin a u nich volí hodnotu nezávisle proměnné. Ve výběrovém šetření nebo jiné observační studii sami hodnocení jedinci volí přiřazení do skupin svým vlastním chováním a výzkumník na tuto volbu nemá vliv. Například při zkoumání vlivu kouření na zdraví by výzkumník rád přiřadil některé experimentální jedince, kteří nejsou kuřáky, do skupiny kuřáků. Avšak toto přiřazení nemůže z etických důvodů provést. Pokud nelze provést experiment, musíme vystačit s observačními studiemi.

Plán nebo schéma experimentu určuje plán sběru dat a řeší s větší či menší jasností problém kontroly rušivých proměnných, jež mohou zkreslit závěry výzkumu. V nejjednodušším experimentu (obr. 2.3) výzkumník vybere jednu skupinu a uplatní na jedincích z této skupiny daný typ intervence nebo ošetření (např. určitý typ tréninku nebo terapie), přičemž hodnotu závisle proměnné měří buď jenom po experimentu (a), nebo i před ním (b). O označuje pozorování (observaci) a X provedení intervence. Pretest, resp. posttest označují první, resp. druhé měření cílové (závislé) proměnné. Tyto dva typy experimentů mají velmi slabou interní validitu. Proto je Campbell a Stanley (1963) nazývají *preexperimenty*.

V komparativním experimentu pracuje výzkumník minimálně se dvěma skupinami. **Kontrolní skupina** obvykle sestává z jednotek, které nejsou vystaveny intervenci, jejíž účinky se zkoumají. Skupina vystavená zkoumanému typu intervence se nazývá **experimentální skupina**. Ostatní podmínky jsou pokud možno u obou skupin udržovány stejně, takže vzniklé rozdílnosti lze přičíst pouze na vrub zkoumané intervence.

Obr. 2.3 Preexperimenty



Experimenty je možné rozšířit na několik ošetřených skupin. V některých komparativních experimentech nemáme kontrolní skupinu bez intervence, ale všechny skupiny jsou ošetřeny nějakým způsobem (třeba i standardní intervencí). V složitějších experimentech měníme záměrně více ovlivňujících nezávisle proměnných. V tomto odstavci se zmíníme o zcela jednoduchých schématech experimentu.

Hlavním problémem při experimentování je navrhnut plán experimentu tak, že jednotlivé ovlivňující proměnné nepůsobí vzájemně rušivě. Přesněji řečeno, dvě proměnné jsou v rušivém vztahu, pokud nemůžeme oddělit (separovat) jejich účinek na závisle proměnnou. Jedním typem potencionální rušivé proměnné je **skrytá proměnná**, kterou nesledujeme v rámci studie. Tato skrytá proměnná může ovlivňovat závisle proměnnou. Samozřejmě v každém experimentu existují skryté proměnné – některé mají rušivý charakter, jiné nikoliv. Například když testujeme nový lék proti vysokému krevnímu tlaku, považujeme za nezávisle proměnnou dávku léku a závisle proměnnou hodnotu systolického tlaku. V tomto kontextu může působit věk, pokud jeho hodnoty v experimentu neuvažujeme, jako skrytá rušivá proměnná. Jestliže jsou na druhé straně jedinci do kontrolní i experimentální skupiny vybráni cestou náhodného přiřazení bez ohledu na věk, je správné vzniklé rozdílnosti tlaku mezi skupinami přičítat na vrub nového léku. V tomto případě říkáme, že náhodné přiřazení jedinců do skupin kontrolovalo působení vlivu věku jako rušivé proměnné. Procedura, kterou dosáhneme podobnosti (homogennosti) porovnávaných skupin, se nazývá **randomizace**.

Určení vztahu mezi závisle proměnnou a nezávisle proměnnými je cílem experimentálního procesu. Nezávisle proměnné nazýváme někdy faktory a jejich hodnoty úrovně faktorů. **Intervence** nebo **osetření** představuje kombinaci specifických hodnot faktorů. Při experimentování uvažujeme obzvláště tyto tři aspekty: komparaci, randomizaci a opakování (replikaci).

Komparace

Experimentální procedury se zaměřují na srovnání. V nejjednodušším případě experimentu provedeme nejdříve u experimentálních jednotek měření uvažovaných proměnných (včetně cílových), aplikujeme posuzovanou intervenci a pak posuzujeme efekt intervence změřením cílových proměnných.

V komparativních experimentech, které používají kontrolní skupinu, je jediná v kontrolní skupině poskytnuta standardní intervence a experimentální skupina je ošetřena novou intervencí. Někdy je kontrolní skupina ošetřena **placebem** – tedy ošetřením, jež je z biologického hlediska zcela neúčinné. Experiment musí počítat s tzv. efektem placebo, jenž způsobuje, že subjekty reagují jakékoli ošetření – i v případě, že se použije placebo. Psychologické působení

bení vědomí, že bylo provedeno ošetření, způsobuje fyziologické nebo psychické změny – podrobně viz Křivohlavý (2001, s. 59–68). Proto se kontrolní skupina ošetřuje placebovou náhražkou, aby bylo možno odhadnout velikost placebového efektu v dané situaci a provést pravdivější srovnání obou skupin.

Po výběru placebo a experimentálního ošetření je nutné přiřadit jednotky k oběma skupinám. Srovnání skupin je možné zajistit, pokud jedinci ve skupinách mají podobné charakteristiky. To lze uskutečnit např. procesem **vyrovnávání** neboli *mačování* (z anglického *matching*), v němž ke každé jednotce v experimentální skupině vyhledáme podobného jedince a zařadíme ho do kontrolní skupiny. To však můžeme provést, pokud máme zjištěny všechny charakteristiky jedince, jež mohou ovlivnit výsledek, a pokud je jedinec s podobnými hodnotami k dispozici. V mnoha případech výzkumník nemůže mít přehled o všech působících rušivých proměnných. Také se stává, že informace o charakteristikách jsou zjišťovány subjektivním způsobem.

zjišťován subjektivními způsobem.

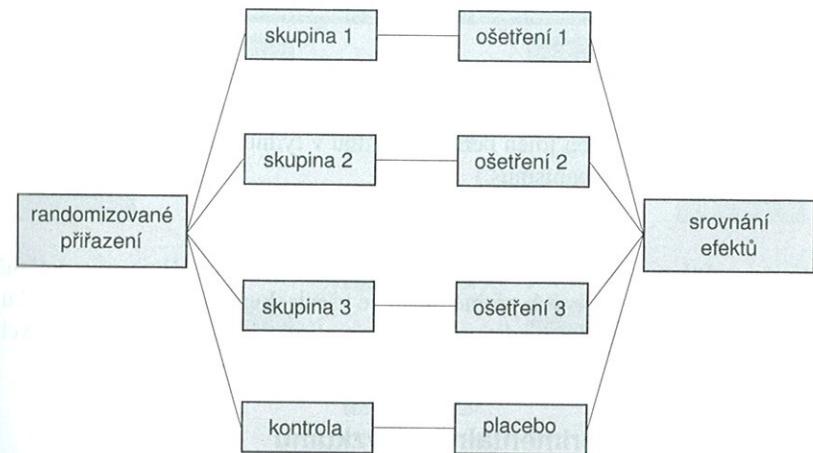
Proto se za účelem získání podobných skupin – např. kontrolní a experimentální – používá metoda randomizovaného přiřazování subjektů do skupin, která využívá pravděpodobnostního mechanismu.

Randomizace, strukturní homogenita

Randomizace (náhodné přiřazení) je proces přiřazování jedinců do experimentální a kontrolní skupiny (skupin může být i více než dvě). Dosahuje se pomocí ní strukturní homogenity skupin, čímž se myslí, že rozložení všech rušivých, resp. matoucích proměnných vázaných na objekty je ve všech skupinách velmi podobné. Obecný plán, jehož jeden příklad je graficky zachycen na obrázku 2.4, se nazývá úplný znáhodněný nebo randomizovaný plán.

Randomizace spočívá v tom, že sledované objekty náhodně přiřadíme do porovnávaných skupin. Provádíme to např. házením přesné kostky, jestliže se dohodnou patřičná přiřazovací pravidla (např. sudé číslo – zařazení do první ze dvou skupin; anebo čísla 5 a 6 – zařazení do třetí skupiny). V této souvislosti se často používají **tabulky náhodných čísel**. Jedná se o seznam náhodných hodnot rovnoměrně rozdělené veličiny v určitém číselném rozmezí a tvaru, což např. znamená, že každé číslo mezi 1 a 999 má stejnou pravděpodobnost, že ho v daném místě seznamu nalezneme (viz tab. I přílohy B). Při praktickém použití této tabulky se musí definovat přiřazovací pravidlo, jež vychází z počtu zkoumaných objektů ve srovnávaných skupinách. Jestliže např. provádíme zařazování očíslovaných $2a$ zvířat (a je přirozené číslo) do dvou skupin, tak odebereme z tabulky tolik různých náhodných čísel, až a z nich je menší nebo rovno $2a$. Tyto čísla pak označují zvířata, která zařadíme do první skupiny, zbytek tvoří druhou skupinu.

Obr. 2.4 Úplný randomizovaný plán



PŘÍKLAD 2.4

Randomizace (náhodné přiřazení) v experimentu

Uvedme příklad náhodného přiřazení do skupiny. Celkem 12 studentů se má zařadit náhodně k jedné ze tří možných „terapií“, každá skupina má obsahovat 4 experimentální jednotky. Studenti se očíslují. Vybere se náhodně dvanáct dvojciferných čísel a ty se přiřadí k označeným studentům (tabulka 2.5).

Tato náhodná čísla dostanou v dalším kroku přiřazena pořadí, která odpovídají jejich velikosti. V našem případě, kdybychom náhodná čísla srovnali podle velikosti, číslo 54 přiřazené studentovi č. 1 by bylo na 7. místě. Studentovi č. 1 přiřadíme proto pořadí 7.

Tab. 2.6 Příklad randomizace – náhodné přiřazení do tří skupin

Študenti	1	2	3	4	5	6	7	8	9	10	11	12
Reálnodné čísla	54	28	06	78	50	96	02	59	61	37	00	76
Pořadí	7	4	3	11	6	12	2	8	9	5	1	10
Skupina	II	I	I	III	II	III	I	II	III	II	I	III

PŘEHLED STATISTICKÝCH METOD

Z těchto pořadí pak vyplýne jednoduché pravidlo pro zařazení studentů do skupin: studenti s pořadím 1–4 jsou zařazeni do první skupiny, s pořadím 5–8 do druhé skupiny a ostatní do třetí skupiny. Tento přístup lze jednoduše přenést i na jiné situace – nestejný počet ve skupinách, libovolný počet experimentálních objektů i skupin apod.

Nelze-li před pokusem označit objekty očíslováním (protože např. jsou pacienti teprve postupně přijímáni), musí se zvolit jiný postup. (Poznamenejme, že výběr podle začátečních písmen jmen nebo podle dnů v týdnu nepředstavuje korektní náhodný výběrový mechanismus.)

• • •
Zatímco právě popsané náhodné přiřazování do skupin rozděluje rovnoměrně hodnoty neznámých rušivých proměnných ve všech skupinách, vytváření bloků, jež popíšeme v následujícím odstavci, dovoluje eliminovat známé vlivy rušivých proměnných.

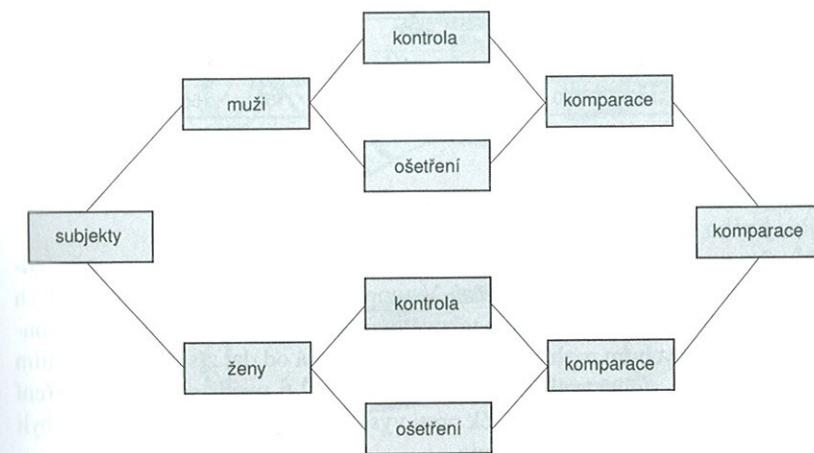
Blokový plán experimentálního výzkumu

Úplný znáhodněný plán patří k nejjednodušším plánům experimentů. Stejně jako při volbě typu výběru, existují i různé plány experimentu, které mohou být v určitých situacích výhodnější. Například vyrovnání (mačování) jedinců různým způsobem může vést k přesnějším výsledkům než uskutečnění úplně znáhodněného experimentu.

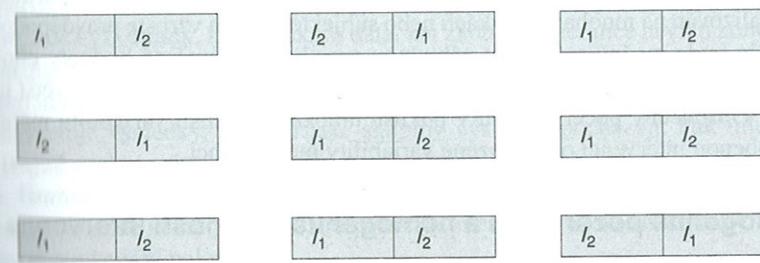
Blok je skupina experimentální jednotek, jež jsou si podobné ve faktoru, o kterém se předpokládá, že ovlivňuje závisle proměnnou. Princip vytváření bloků odpovídá získání stratifikovaného výběru využitím známých rušivých proměnných jako věk, pohlaví atd. jako stratifikačního kritéria. Uvnitř bloků následuje randomizace do skupin. Tím se dosáhne toho, že variabilita způsobená rušivým faktorem je uvnitř bloků menší než mezi bloky. Při vyhodnocování získaných dat se schéma tvorby bloků bere v úvahu: např. bloky se vzájemně statisticky porovnají.

Například při experimentálním testování léku proti vysokému tlaku může hrát roli pohlaví osoby. Na obrázku 2.5 je graficky zachycené blokové upořádání, které dovoluje cíleně rozlišit variaci výsledků v důsledku působení znaku pohlaví od efektů ošetření. V tomto upořádání je každá kategorie pohlaví považována za od sebe nezávislou jednotku. Je důležité, aby se uvnitř bloku provedlo zařazení do kontrolních a ošetřovacích skupin náhodně. Blokové uspořádání je podobné stratifikovanému výběru, jejž používáme u statistického šetření. V obou případech strata nebo bloky jsou tvořeny se záměrem kontrolovat efekt známých rušivých proměnných.

Obr. 2.5 Blokové uspořádání podle ošetření a pohlaví



Obr. 2.6 Porovnání účinků intervencí I_1 a I_2 pomocí vyrovnání (mačování)



Jinou blokovou strategií je plán s vyrovnanými (mačovanými) páry. V tomto plánu kontrolní a ošetřené skupiny vytváříme párováně pro stejné hodnoty ovlivňujících proměnných, které můžeme nebo nemusíme znát předem. Příkladem pro vyrovnaní může být porovnání dvou různých intervencí graficky znázorněné na obrázku 2.6. V tomto pokusu každá z 9 škol představuje blok, ve kterém se provádí dvě třídy, k nimž náhodně přiřadíme danou intervenci. Jednotlivé školy zajišťují výběr z různých typů škol. Při experimentování se subjekty se v ta-

PŘEHLED STATISTICKÝCH METOD

Tab. 2.6 Příklad spárovaných měření – srovnání krevního tlaku před terapií a po ní u stejných pacientů

Jedinci	1	2	3	4	5	6	7	8	9
Tlak před terapií	150	165	145	175	156	180	195	152	160
Tlak po terapii	140	170	135	165	150	152	156	160	143

kových pokusech někdy používají jednovaječná dvojčata. Také existují přirozené dvojice jiného původu, např. manželské páry.

Mnoho studií se opírá o měření objektů před pokusem a po pokusu, což představuje jeden ze způsobů vyrovnavání. Vstupní test a závěrečný test na stejných žácích ve škole představuje příklad takového přístupu. Někdy musíme znát kontext výzkumu, abychom rozlišili takto spárovaná data od dat získaných měřením dvou nezávislých stejně velkých skupin. Tabulka 2.6 uvádí spárovaná měření experimentu, v kterém se testoval lék proti vysokému tlaku. Vybraní pacienti byli změřeni na začátku a na konci terapie.

Opakování

Opakováním se myslí uskutečnění stejného pokusu na mnoha jednotkách. To pomáhá snížit roli náhody při hodnocení výsledků. Je nutné, aby experiment byl realizován na mnoha jednotkách nebo subjektech. Tím vzrosté pravděpodobnost, že bude možné redukovat náhodnou variabilitu průměrné hodnoty cílové proměnné v jednotlivých skupinách a identifikovat tak efekt ošetření s větší jistotou. Dostatečný počet jedinců v pokusu umožňuje rozlišit variabilitu měření totou. Dostatečný počet jedinců v pokusu umožňuje rozlišit variabilitu měření totou. Dostatečný počet jedinců v pokusu umožňuje rozlišit variabilitu mezi jedinci. způsobenou intervencí od přirozené variability mezi jedinci.

Homogenita pozorování a homogenita okolností intervence

Homogenitou pozorování rozumíme okolnost, že měřené znaky se na zkoumaných jednotkách zachycují jednotným způsobem u všech porovnávaných skupin. Proto se určují přesné postupy měření, laboratorní postupy a vyloučují se subjektivní vlivy na pozorování jak u hodnotitele, tak u hodnoceného.

Vyloučení subjektivního vlivu na pozorování se dosahuje slepými pokusy neboli maskováním či zaslepením. Při jednoduchém slepém pokusu se měřeny jedinec nedozvídí, jaká forma ošetření je u něho prováděna. Měl by však být rámcové informován o principu slepého pokusu. **Jednoduchý slepý pokus** je vhodný tehdyn, když data pocházejí ze subjektivních pozorování subjektů (test, celkový pocit apod.). Jestliže existuje možnost, že i výzkumník (hodnotitel)



„Pane profesore, jste si jistý, že takto vypadá dvojitě slepý pokus?“

podléhá silným subjektivním vlivům, např. rád by upřednostnil některé ošetření, pak tento dojem může přenést i na hodnoceného a tím ovlivnit jeho posuzování výsledků. Tomuto kombinovanému jevu se předchází **dvojitým slepým pokusem**, kde výzkumník, který získává data, ani zkoumaný jedince neví, u koho bylo použito hodnocené ošetření. Je zřejmé, že slepé pokusy nejsou vždy praktické pro mnohé otázky ani nutné. Další způsob maskování znamená, že ani ten, kdo interpretuje výsledky, neví od jaké skupiny výsledky pocházejí. Pak mluvíme o **trojitě zaslepeném** nebo **trojitě maskovaném** porovnávání.

Homogenita okolností intervence vyjadřuje požadavek, že srovnávané skupiny se od sebe neliší podmínkami průběhu intervence až na znak, který určuje důvod pro konání pokusu.

Další plány experimentů

Faktorový plán uvažuje více než jeden faktor. Zahrnuje vytvoření skupin pro všechny možné kombinace úrovní všech faktorů, takže relativně jednoduchý faktorový plán např. s třemi faktory a dvěma úrovněmi u každého faktoru musí vytvořit 8 nezávislých skupin s náhodným přiřazením. Omezíme se na faktorový plán s dvěma faktory I a II o dvou úrovních 1 a 2 s měřením po provedení intervencí. Postupujeme následovně:

PŘEHLED STATISTICKÝCH METOD

1. Vytvoříme čtyři experimentální skupiny použitím náhodného přiřazení.
2. Experimentální skupina A bude ošetřena tak, že oba faktory mají úroveň 1.
- Experimentální skupina B bude ošetřena tak, že faktor I má úroveň 1 a faktor II má úroveň 2.
- Experimentální skupina C bude ošetřena tak, že faktor I má úroveň 2 a faktor II má úroveň 1.
- Experimentální skupina D bude ošetřena tak, že oba faktory mají úroveň 2.
3. Provedeme posttest u všech skupin.

Tento plán nám umožňuje zjistit efekty odděleně pro každý faktor a také, zda tyto „hlavní“ efekty pro daný faktor závisí na úrovni druhého faktoru. Tento postup je možné doplnit o provedení pretestu.

Plán s dvěma skupinami, pretestem a posttestem pracuje s experimentální a kontrolní skupinou a zárazuje pretest. Použití pretestu má výhody i nevýhody. Výhody jsou zřejmé. Můžeme jím překontrolovat randomizaci. Jestliže skupiny nejsou shodné v pretestu, můžeme při hodnocení upravit výsledky posttestu pomocí vhodné statistické metody (analýza kovariance, kap. 10.7). Diference mezi skupinami mohou být vlivem intervence. Jedná se o citlivější vyhodnocení než pouhé srovnání posttestů. Hlavní nevýhodou je senzitivizace pretestem, která může ovlivnit posttest. Při tomto experimentu se postupuje takto:

1. Vytvoříme kontrolní a experimentální skupinu náhodným přiřazením.
2. Provedeme pretest u obou skupin.
3. Experimentální skupinu podrobíme intervenci. Kontrolní skupina dostane „placebové ošetření“.
4. Provedeme posttest u obou skupin.

Protože randomizované kontrolované pokusy jsou velmi nákladné, používají se někdy tzv. **historické** kontrolní skupiny. Přitom se srovnávají výsledky ošetřené skupiny se skupinou výsledků, které se získaly z historických záznamů o jednotlivých, jež nebyly ošetřeny nebo dostaly standardní ošetření. Hlavní nevýhodou je potíž s dosažením srovnatelnosti ošetřené a historické kontrolní skupiny ve všech aspektech. S podobnými obtížemi se setkáváme u tzv. **kvaziexperimentů**, kdy tvoříme komparované skupiny bez uplatnění principu randomizace.

Plán s křížovou klasifikací a opakováním měření. Validita plánů s opakováním měření je ohrožena tím, že stejní jedinci jsou vystaveni experimentování a měření několikrát. To vyvolává tzv. efekt pořadí: jedinci jsou ovlivněni předchozím ošetřením a měřením (učí se reagovat na test, jsou unaveni, atd.). Tento efekt lze minimalizovat pečlivým organizováním celého experimentu. Řešení

také představuje křížová klasifikace. To znamená, že jedinci jsou náhodně přiřazeni k různým pořadím různých ošetření. Popíšeme jednoduchý plán s opakováním měření s křížovou klasifikací:

1. Vytvoříme dvě skupiny pomocí náhodného přiřazení.
2. Skupina A se podrobí nejdříve ošetření 1 a pak ošetření 2. Skupina B se podrobí nejdříve ošetření 2 a pak ošetření 1.
3. Obě skupiny jsou měřeny po každém ošetření.

Parametrický plán znamená vytvoření skupin s několika hladinami nezávisle proměnné. Popíšeme jednoduchý parametrický plán s jedním faktorem, který samozřejmě může zahrnout i komparativní skupinou s „nulovým“ ošetřením.

1. Vytvoříme tři experimentální skupiny použitím náhodného přiřazení.
2. Experimentální skupina A dostane ošetření na úrovni 1 sledovaného faktoru. Experimentální skupina B dostane ošetření na úrovni 2 sledovaného faktoru. Experimentální skupina C dostane ošetření na úrovni 3 sledovaného faktoru.
3. Provedeme posttest u všech tří skupin.

Validita závěrů experimentálních studií

Splnění požadavků na interní a externí validitu je předpokladem hodnověrnosti studie. Při plánování výzkumu se snažíme o maximalizaci obou validit. Zpráva o výzkumu má obsahovat v diskusi poukazy na možnost porušení interní validity a důsledky omezení externí validity.

Interní validita je předpokladem pro hodnověrnost závěrů z výsledků experimentu. Vyjadřuje stupeň, do jaké míry plán experimentu nedovoluje jiné vysvětlení změn než to, které vychází ze zkoumané hypotézy. Campbell a Stanley (1963) identifikovali osm možných mechanismů porušení interní validity. Některé z nich již zmínili. Zde provedeme jejich souhrnný přehled. Porušení interní validity spojuje, že danou konfiguraci výsledků, jež svědčí pro nebo proti hypotéze, lze vysvětlit působením jiného mechanismu než na základě zkoumané hypotézy. Pokud se nežádoucí mechanismy mohou při zkoumání vyskytnout, musíme provést před experimentem opravu plánu nebo přikročit po experimentu ke zkoumání významnosti jejich efektů, modifikaci analýzy a k příslušné korekci našich závěrů.

Za účelem demonstrace předpokládejme v dalším výkladu, že zkoumáme nového způsobu výuky. Navrhne studii, v níž srovnáváme novou metodu s tradičním postupem. Co se vše může stát a vést ke zkreslení výsledků?

PŘEHLED STATISTICKÝCH METOD

1. **Selekce.** Můžeme vybrat děti tak, že se ve skupinách ocitnou jedinci s nestejnými charakteristikami, které ovlivňují učení. Jedna skupina může být „chytřejší“ nebo má větší zkušenosti s probíranou látkou.
2. **Maturace.** Během studie může jedna skupina prodělat změny bez vztahu k programu (tělesné, emoční) rychleji (pomaleji) než druhá. Tento falešný efekt programu zůstane nejčastěji nezpozorován, pokud zkoumáme změny jenom u jedné skupiny.
3. **Historie.** Během studie se může stát něco jiného, co nemá v vlastním programu nic společného a ovlivní to jednu ze skupin. Například na školu přijde lepší učitel nebo komfortnější vybavení, v místnosti je z důvodu závady zima nebo teplo. Čtenář by jistě sám vymyslel bezpočet věcí, jež se mohou stát a ovlivnit výsledek.
4. **Mortalita** neznamená v tomto případě doslova úmrtnost, ale jakékoli ztráty. Z nějakých příčin jednu skupinu opustí více lidí. Pokud to budou žáci, kteří nejsou s programem spokojeni, pravděpodobně to zlepší u této skupiny její konečné průměrné výsledky.
5. **Regres k průměru.** Na základě předchozích testů jsme zařadili do jedné skupiny žáky s lepšími výsledky, je pravděpodobné, že na konci budou mít výsledky relativně horší než na začátku, aniž by to bylo důsledkem působení programu. Stalo se tak působení statistického jevu „regrese k průměru“, který popisujeme v kapitole 7.4.
6. **Testování.** Při používání stejného testu na začátku a v průběhu programu dáváme žákům možnost si na test zvyknout a dosahovat v něm lepší výsledky.
7. **Použité měřicí postupy.** Výsledky testu jsou nespolehlivé, protože test byl špatně proveden. Procedura měření se během pokusu změnila.

Provedení randomizace a homogenita provedení pokusu je odpovědí na většinu potíží s interní validitou. Také je nutné pracovat s většími rozsahy výběrů, aby randomizace fungovala bezpečně. V takto provedených pokusech pak můžeme předpokládat, že všechny faktory ovlivňující výsledky jsou v obou skupinách stejně rozloženy. Pokud není randomizace možná, vyplatí se zjistit hodnoty dostatečně rušivých faktorů u obou skupin. Nemůžeme sice zabránit jejich působení, ale při analýze lze údaje o nich vzít často v úvahu a provést korekci výsledků např. pomocí analýzy kovariance (kap. 10.7).

Působení rušivých proměnných můžeme také předcházet pečlivým tréninkem personálu, jenž se na pokusu podílí. Data získaná monitorováním průběhu pokusu se použijí při kontrole validity získaných závěrů. Čemu nelze zabránit přípravou pokusu a školením personálu i účastníků studie, měli bychom aspoň pozorovat a zaznamenávat, abychom příslušná data pak využili při analýze celého pokusu.

Zatímco dosud popisovaná interní validita je koncept zajišťující interpretovatelnost výsledků, **externí validita** se zabývá otázkou přenositelnosti výsledků na jiné situace, zobecnitelností závěrů pokusu na jiné případy. Při realizaci pokusu se snažíme maximalizovat interní validitu a kontrolovat všechny faktory, jež ji mohou porušit. Tato snaha má však ten efekt, že pracujeme v umělých podmínkách, které v běžné praxi školní výuky nebo ve zdravotnickém prostředí nikdy nemohou nastat. Platí pak naše závěry ze studie i pro tyto „reálné“ situace? Externí validita zkoumání je stupeň, do jaké míry výsledky našich závěrů zůstanou v platnosti při změněných podmínkách.

Existují dvě kategorie ovlivnění externí validity. Jedna se nazývá populační validita, druhá ekologická validita. **Populační validita** se zajímá o pokusné jednotky ve výzkumu. Jestliže změnите věk, pohlaví, schopnosti jedince, změnil by se výsledek pokusu? **Ekologická validita** se zajímá o ostatní stránky pokusu: kdo dělal co, kde, jak a kdy. Kdyby se studie konala v jiném prostředí (např. mimo laboratoř), kdyby se použil jiný test a pokus trval kratší dobu, byly by výsledky podobné?

Externí validita se liší od interní validity v několika směrech:

1. Experiment může být interně validní, ale nemusí mít vysokou externí validitu. Na druhé straně, pokud pokus není interně validní, nemůžeme mluvit ani o jeho externí validitě. Studie s jasnými chybami nemohou vést k nějakému zobecnění.
2. Interní validita pojednává nerovnosti mezi skupinami v pokusu. Externí validita pojednává pokus v celku, nevšímá si jednotlivých skupin nebo jedinců.
3. Aspekty externí validity se týkají všech výzkumů. Bez ohledu na použitou metodu můžeme uplatnit na výsledky hledisko externí validity.

Víme-li si podrobněji aspektů, jež vedou k ohrožení externí validity.

I. Faktory týkající se populační validity:

- a) **Populace v experimentu, cílová populace.** Tyto dvě množiny se mohou značně lišit. Můžeme přenést výsledky z pokusu s desetiletými školáky na všechny školáky?
- b) **Interakce mezi ošetřením (programem, terapií) a osobnostními charakteristikami.** Jedinci v pokusu mohou mít nějaké zvláštní vlastnosti, které v interakci s ošetřením mohou ovlivnit výsledek studie. Pokus je interně validní pro získaný vzorek, ale budou validní v případě, že místo dětí z města budeme pracovat s venkovskými dětmi, kteří na vyučovací metodou reagují jinak?

2 Faktory týkající se ekologické validity:

- Faktory týkající se ekologické vandy.

 - a) **Definice konstruktů, závisle a nezávisle proměnných.** Obvykle výzkumníci použijí specifické definice konstruktů a proměnných, což ovlivní další průběh výzkumu a zobecnitelnost získaných výsledků. Žádný výzkum nemůže obsáhnout všechny možnosti pojímání určitých pojmu. Vzniká tak situace, že používáme podobná slova, ale máme na mysli něco jiného. Je použitá definice školní úspěšnosti v posuzovaném výzkumu relevantní pro můj výzkum nebo situaci?
 - b) **Interference více ošetření.** Ve studii se může použít více ošetření jednoho jedince najednou nebo jedno ošetření může mít více částí. Interagují tyto části mezi sebou, aby pak vznikl kumulativní efekt, nebo každá část nebo ošetření má svůj vlastní efekt bez ohledu na uplatnění ostatních částí ošetření? Jestliže zopakujeme pokus pouze s jedním ošetřením, bude výsledek podobný?
 - c) **Hawthornský efekt, efekt novosti, efekt rivalry, efekt očekávání experimentátora.** Tyto efekty nastávají obvykle u špatně navržených a provedených výzkumů. Subjekty vědí, že jsou účastníci výzkumu, takže se chovají jinak než v běžném životě. Ošetření je tak neobvyklé, že je pro účastníky zajímavé a vzrušující. Tak vzniknou výsledky, jež by po uvedení programu do praxe již nebyly tak průkazné. Efekt rivalry vznikne, když se skupina domnívá, že je o něco ošizena (je ošetřována „horším“ programem) a domnělé poškození se snaží nahradit zvýšeným úsilím. Efekt experimentátora se projevuje tak, že účastníci pokusu se snaží přizpůsobit očekávání experimentátora, který dává některému z ošetření přednost a očekává určitý typ chování.
 - d) **Senzitivizace mezi pre- a posttestem.** Efekty vznikly, protože jsme jeměřili. Kdybychom měření vůbec neprováděli, výsledky by byly jiné.
 - e) **Interakce mezi historií a ošetřením.** Výsledky mohou být způsobeny nějakým vnějším jevem, jenž nastal ve stejnou dobu, kdy probíhal pokus. Zvýšily se příspěvky pro Červený kříž v důsledku reklamy v televizi, anebo proto, že byly ve stejnou dobu hlášeny velké záplavy?
 - f) **Interakce mezi dobou měření a ošetřením.** Získané výsledky se záležaly měřením a jejich hodnota je závislá na době měření, ne na efektu programu. Určité měření lze provést na začátku nebo na konci školního roku, bude mít okamžik pokusu vliv na výsledky?

Je nutné si uvědomit, že žádný pokus nemá ideální externí validitu. Vždycky existují pochybnosti. V této souvislosti se nejčastěji kritizují výzkumy prováděné

v umělých podmínkách laboratoře. Jednou z cest pro zvýšení externí validity představuje zkoumání situací v přirozených podmínkách pomocí případových studií. Mluvíme pak někdy o tzv. naturalistickém výzkumu.

Shrnutí důležitých aspektů experimentování

Podstatným znakem experimentu je, že výzkumník záměrně volí různé podmínky (různé hodnoty nezávisle proměnných, různá ošetření) a sleduje, jak ovlivňují cílovou sledovanou proměnnou. Při provedení experimentu se snažíme zajistit zejména tři vlastnosti – existenci kontrolní skupiny, randomizaci a opakování. Pokud jsou objekty rozdeleny do ošetřených a kontrolních skupin náhodně, lze odhadnout, jaké velké rozdíly mezi skupinami mohou být způsobeny pouze přirozenou náhodnou variací. Je-li rozdíl zjištěný mezi skupinami větší než tato přirozená variace, mluvíme o statisticky významném výsledku, jejž při čítáme působení ošetření.

V určitých případech lze po provedení experimentu, vyhodnocení všech vlivů snižujících interní validitu studie a zjištění rozdílných účinků prohlásit, že jsme ověřili základní podmínky kauzálního vztahu mezi nezávisle proměnnou (faktor ovlivnění) a závisle proměnnou. V procesu posuzování kauzality vztahu ověřujeme platnost dalších kritérií (viz také s. 58).

Uvádíme kontrolní otázky, které musíme zodpovědět při navrhování experimentu i při kritickém posuzování získaných výsledků:

- I Kdo budou účastníci studie? Do jaké patří populace?
 - I Jak budou vybráni? Bude použit náhodný výběr?
 - I Použije se náhodné přiřazení? Jak bude provedeno?
 - I Kolik bude jedinců v každé skupině? Použije se pro určení počtu nějaký statistický výpočet?
 - I Co je závislá (cílová) proměnná ve studii? Kolikrát bude měřena?
 - I Co je ošetření? Jak je operacionalizované?
 - I Budou měřeny nějaké kovarianční (doprovodné) proměnné?
 - I Jak bude vypadat plán experimentu? Jak vypadá jeho grafické zachycení?
 - I Jaké měřící instrumenty se použijí? Kdo je navrhl? Proč byly vybrány? Byla ověřena jejich spolehlivost a validita?
 - I Jaké jsou kroky celého postupu (např. náhodné přiřazení, zjištění demografických údajů, měření pretestu, provedení ošetření, měření posttestu)?
 - I Jaké jsou potencionální překážky pro zajištění interní a externí validity? Jak je překonají?
 - I provede se pilotní studie?
 - I Jaké se použijí statistické metody pro vyhodnocení dat (popisné a inferenční)?

2.4.4 Kategorizace výzkumných plánů

Existuje několik kategorizačních systémů výzkumných plánů a výzkumných strategií. Jednoduchou klasifikaci používá ve své příručce metodologie pro menší výzkumné projekty Robson (1993). Ten považuje za tradiční tři výzkumné strategie: experiment, výběrové šetření (survey) a případovou studii (tab. 2.7). Robson pětadvacet let po svém vydání uvedl, že při obecném pohledu lze všechny předchozí strategie zahrnout pod pojmem případové studie, protože každá studie je omezená v prostoru a času a odehrává se v daném kontextu. Jednotlivé studie představují epizody vedeckého výzkumu. Tuto myšlenku interpretuje po svém metaanalytický přístup, jejž popisujeme v kapitole 14. Více o případové studii čtenář najde v příslušné literatuře (např. Yin, 1989; Hendl, 1999).

Velmi známý kategorizační systém navrhli metodologové Campbell a Stanley (1963). Popisuje ho podrobně Ferjenčík (2000). Podle Campbella a Stanleye jde o to, jak pomocí výzkumného plánu získat data, která umožní odhadnout populační charakteristiky nebo efekty intervencí bez systematického zkreslení a s dostatečnou přesností. Tyto požadavky představují analogii k požadavkům na kvalitu měřicích metod. Každý plán má v tomto smyslu své silné a slabé stránky.

Tabulka 2.8 popisuje v přehledu základní kategorie výzkumných plánů, jež se uplatňují v empirickém kvantitativním výzkumu.

Tab. 2.7 Klasifikace výzkumných plánů podle Robsona (1993)

Výzkumný plán	Podstata	Typický postup
Experiment	Měří se efekty, které vznikají, když se manipuluje s jednou nebo více nezávislými proměnnými.	Vybere se vzorek populace, jedinci (nebo objekty) se přiřadí k různým experimentálním podmínkám, uskuteční se plánované změny nezávislých proměnných. Měří se malý počet cílových proměnných. Jiné proměnné se mohou zohlednit při analýze.
Výběrové šetření	Sběr informací ve standardizované podobě od skupiny lidí.	Vybere se vzorek ze známé populace, od každého prvku vzorku se získá určité množství informací standardizovaným způsobem (obyčejně pomocí dotazníku nebo pomocí strukturovaného interview). Analýza dat.
Případová studie	Získání podrobné, hloubkové znalosti o jednom případu nebo o několika málo „případech“.	Vybere se případ, situace, jedinec, instituce nebo skupina, studuje se v daném kontextu. Pro sběr dat se užívají různých metod (pozorování, interview, analýza dokumentace). Analýza dat.

Tab. 2.8 Kategorie výzkumných plánů

Typ výzkumného plánu	Účel	Míra kontroly rušivých proměnných	Charakteristika
1. pravý experiment (např. pretest-posttest)	srovnání různých skupin (ošetřené a kontrolní) s randomizací	vysoká míra záměrné manipulace	manipulace nezávisle proměnnou a náhodně přiřazení
2. kvaziexperiment	srovnání skupin, vnitroskupinový experiment	určitá míra záměrné manipulace s proměnnými	manipulace nezávisle proměnné, bez náhodného přiřazení, často více měření v čase
3. preexperiment (např. pouze posttest)	srovnání skupin nebo hodnocení skupiny, většinou bez sledování v čase	minimální míra záměrnosti a určení hodnot proměnných	manipulace s nezávisle proměnnou nebo měření proměnné v jedné skupině bez možnosti kontroly a náhodného přiřazení
4. popisný výzkum (census, výběrové šetření, observační studie, korelační a komparativní studie)	popis události nebo fenoménu	bez možnosti záměrného určení působení rušivých proměnných	bez manipulace s nezávisle proměnnými, bez náhodného přiřazení

Klasifikace výzkumných plánů se odlišují mezi obory. V medicíně se uplatňují podobná kritéria, ale jiná terminologie pro klasifikaci a popis výzkumných plánů. Například klinická epidemiologie při výzkumu rizikových faktorů nebo efektivity intervencí rozlišuje:

- 1. **Studie případové a série případů.** Podrobně se popisuje případ nebo několik případů.
- 2. **Výběrová statistická šetření – průřezová šetření.**
- 3. **Studie případů a kontrol (retrospektivní příčinně-komparativní výzkumy).** V studii případů i kontrol jsou vybrány osoby s určitou nemocí a bez ní. Poté se sbírájí zpětně údaje o vystavení rizikovým faktorům a usuzuje se o jejich kauzální souvislosti s nemocí.
- 4. **Kohortové studie – longitudinální prospektivní.** V kohortové studii jsou dvě nebo více skupin osob vybrány na základě odlišností v určité charakteristice (např. přítomnost nějakého rizikového faktoru) a pak jsou dlouhodobě sledovány.
- 5. **Randomizované klinické studie – pravé experimenty.**
- 6. **Metaanalýzy a systematické přehledy.** Jde o popisné retrospektivní zpracování výsledků množiny klinických studií.
- 7. **Hierarchie plánů výzkumu z hlediska validity závěrů vzhledem k průkazu příčnosti stojí nejvyšše randomizované klinické studie a metaanalytické studie (5 a 6), které hodnotí všechny specifikované (nejlépe randomizované) primární studie. Průkaznost spíše signálního charakteru ne případová studie (1).**

PŘEHLED STATISTICKÝCH METOD

Poznamenejme, že v jedné výzkumné akci lze použít více výzkumných metod a plánů.

2.5 Organizace dat a jejich kontrola, scházející údaje

Data shromažďujeme v různých kontextech. Často se nejdříve zapisují do speciálních formulářů, např. dotazníku nebo laboratorního formuláře pro zaznamenání údajů. Při přímém elektronickém sběru se údaje ukládají rovnou do počítače. Když jsou všechna data nashromážděna, připravujeme z nich **tabulku dat** nebo **datovou matici**. Datová matice uspořádává přehledně data tak, že v každém jejím řádku je popis jednoho objektu a každý sloupec obsahuje data pro jednu proměnnou. Průsečík řádky a sloupce se označuje jako *buňka* a obsahuje právě jeden údaj nebo hodnotu. Zároveň se ptáme, zda objekty jsou nějakým způsobem třídeny do skupin, které musíme vzít při analýze v úvahu. Pokud jsme tak již neudělali, snažíme se navrhnut vzhodné klasifikační proměnné a doplnit jejich hodnoty.

Jako příklad uvádíme část datové matice z výzkumu úrovně pohybových dovedností vybraných třináctiletých školáků a školaček, u nichž máme navíc poznamenáno jejich pohlaví, údaj o organizovanosti ve sportovním oddíle a údaj o prospěchu z matematiky (tab. 2.9).

Kódování dat

Data mohou být různým způsobem kódována. Kód znamená jednoznačný předpis, jak přiřazovat určitým hodnotám proměnných vhodné symboly, většinou čísla. Kódování může vycházet např. z požadavků počítačového programu, který budeme používat pro zpracování. Pro začátečníky navrhujeme dodržovat několik pravidel, jež jsou podložena zkušenosťí:

1. Pro všechny proměnné musí být jasné vymezen způsob kódování. Například při kódování pohlaví použijeme přiřazení: muž – 0, žena – 1.
2. Každá proměnná musí zaujmít pro každý objekt stejnou pozici. To znamená, že v tabulce je jí přiřazen jediný sloupec.
3. Všechny kódy pro proměnnou musí být disjunktní a vyčerpávat všechny její možné hodnoty. To znamená, že daná proměnná nemůže nabývat zároveň dvou různých hodnot.
4. Kódování má zachovat maximum informací pro proměnnou. Vyhýbáme se umělé redukci kvantitativních hodnot na dichotomické apod. Například při kódování věku nepřiřazujeme přesně udanému věku kategorizovanou hodnotu podle dekád.

Tab. 2.9 Příklad datové matice (tabulky dat) – údaje o studentech

Jedinec	Skok do délky [m]	Sprint 75 m [s]	Pohlaví	Sportovní organizovanost	Prospěch z matematiky	Skok v závodu [m]
1	3,3	14,3	0	0	3	3,15
2	3,4	13,7	1	0	2	3,5
3	3,35	14,1	1	1	3	3,45
4	3,75	12,3	0	1	1	3,7
5	3,15	13	1	0	1	3,25
6	3,3	12	1	0	2	3,4
7	3,45	13,3	0	0	2	3,55
8	3,5	13,2	1	1	3	3,35
9	3,65	11,1	0	0	4	3,5
10	4,05	10,9	0	1	2	4,15
11	3,35	13,1	1	0	2	3,45
12	3,5	13,1	0	0	4	3,15
13	3,75	12,2	0	0	4	3,95
14	3,55	12,5	1	1	3	3,45
15	4,2	10,8	0	1	2	4,15
16	3,5	12,2	1	0	2	3,75
17	3,7	13,6	1	1	1	3,65
18	3,15	13,4	1	0	1	3,25
19	3,9	11,8	0	1	3	3,8
20	3,65	11,7	0	1	2	3,55
21	4,05	10	0	1	1	4,1
22	3,65	12,4	0	0	3	3,55
23	3,1	13,9	1	0	3	3,2
24	3,6	11,9	0	0	2	3,5
25	3,9	12,5	1	1	2	3,95
26	3,65	12,8	1	0	2	3,6

U proměnné „pohlaví“ užíváme kódování muž – 0, žena – 1;

U proměnné sportovní organizovanost kódujeme „ano“ – 1; „ne“ – 0.

Musíme definovat kódy pro každou proměnnou a pro všechny její specifické hodnoty. Například je nutné kódovat i scházející údaje zvláštními kódy, které musí pokrýt varianty jako: osoba odmítla odpovědět; osoba není doma; během dotazování po telefonu osoba zavěsila; tazatel přeskочil několik otázek; nahývací zařízení přestalo fungovat apod.

Musíme se vyrovnat s kódování nespecifických odpovědí (nevím, nejsem zahrnut) a s rozlišením mezery a nuly.

Kontrola dat

Garbage in, garbage out (GIGO) – smetí dovnitř, smetí ven je staré přísloví, strašící všechny výzkumníky. Výsledky statistické analýzy jsou tak dobré, jak dobrá jsou data.

Při zkoumání chyb, jimiž mohou být data zatížena, je vhodné odlišovat chybu měření a chybu zpracování. Chybami měření jsme se již zabývali v kapitolách 2.1 a 2.2. **Chyby zpracování** nastávají při organizaci dat, při přenosu získaných dat do datové tabulky a do počítače. Tyto chyby se mohou objevit na jakékoli úrovni zpracování dat. Příklady tvoří různé špatné posuny desetinných čárek, chyby při přepisování – záměna znaků (místo číslice „0“ se objeví znak pro písmeno „O“) atd. Základní strategie pro minimalizaci výskytu chyb je určení jednotlivých fází, kdy se mohou objevit, a pokus jim předejít nebo je odstranit. To zahrnuje zvýšenou pečlivost při záznamu dat, ruční a počítačovou kontrolu nebo duplicitní způsob vstupování dat. Také se uplatňují statistické procedury pro hledání odlehlých hodnot, kdy určujeme logicky nebo statisticky netypické hodnoty v množině měření, které dále přezkušujeme, opravujeme nebo vylučujeme. Kontrola kvality dat je samostatná a rozsáhlá problematika. Uvedeme jen několik poznámek:

- Poté, co jsme data vpravili do počítače, je vhodné je vytisknout, a rádek po rádku překontrolovat.
- Pro snadnější odhalení odlehlých a podezřelých hodnot vytváříme předběžné grafické zobrazení dat. Jedním grafem zobrazíme najednou stovky údajů s dostatečnou rozlišovací schopností. Grafickou analýzu provádíme především u větších množin dat. K tomu používáme histogram, krabicový graf, graf lodyhy a listů nebo dvojrozměrné bodové grafy (viz dále).
- Počítač umožňuje uskutečnit různé formy logické kontroly dat.
- Někdy uvádíme v tabulce sloupec, jenž je odvozen z dvou jiných sloupců matice dat, aniž si uvědomíme, že chybu zpracování zmenšíme, když se tento sloupec vypočte počítačovou procedurou.
- Při vyznačování scházejících hodnot nikdy nepoužíváme „0“; vždy volíme vhodnější kód, na který upozorníme své spolupracovníky.

Všechna data obsahují chyby (hlavně když si myslíme, že tomu tak není). Naštěstí existují techniky analýzy dat, které jsou vůči odlehlým hodnotám rezistentní, resp. necitlivé. To vede k paradoxnímu závěru: někdy potřebujeme techniky citlivé na odlehlé údaje (při jejich odhalování), jindy zase potřebujeme techniky rezistentní vůči odlehlým hodnotám (jestliže nemáme čas na jejich odhalování a chceme dojít k nějakým spolehlivým závěrům).

Scházející hodnoty

Je skoro jisté, že se některé údaje pro daný objekt nepodaří získat. Takové údaje identifikujeme jako „scházející hodnota“. Představují nepříjemnou realitu každého výzkumu. Jaké má výzkumník možnosti? Má tyto hodnoty ignorovat nebo je nahradit nějakou přípustnou hodnotou, aby byly všechny záznamy kompletní? Je důležité zjistit, zda scházející hodnoty jsou situovány náhodně. Pokud je jich hodně v určité proměnné, je možné, že subjekty měly potíže při vyplňování dané položky. V některých případech položka *nemá* být vyplněna. Například muž nevyplňuje, kolik měl porodů. Nejdříve výzkumník oceňuje rozsah množství scházejících hodnot u daných proměnných nebo v subpopulacích. Poté se snaží zjistit důvod k nevyplnění.

Jedním ze způsobů, jak zjistit, zda scházející hodnoty jsou v záznamech zastoupeny náhodně, je rozdelení vzorku na skupinu záznamů bez scházejících hodnot pro danou proměnnou a skupinu se scházejícími hodnotami. Pak porovnáváme ostatní charakteristiky obou souborů. Například v jednom z nich je převaha mužů. Nebo uměle kódujeme scházející údaj jako 0 a vyplněný údaj jako 1 a korelujeme proměnné mezi sebou. Tím odhalíme, zda scházející údaje v jedné položce nejsou svázány s jinou nevyplněnou položkou.

Jakmile výzkumník získal přehled, proč nebyly položky vyplněny, rozhoduje pro další kroky. Jestliže chce nebo musí pracovat pouze s úplnými záznamy, může záznamy se scházejícími údaji prostě vymazat. To se provádí pouze tehdy, když scházející údaje neovlivní systematicky výsledky zpracování. Připravujeme tak o informace. Také můžeme vyřadit ze zpracování pouze celou proměnnou, v níž je příliš scházejících údajů.

Druhým hlavním přístup spočívá v doplnění scházejících hodnot. Tomuto procesu se říká **imputace**. Imputace se provádí několika způsoby. Například odhadujeme scházející hodnoty pomocí regresní analýzy v závislosti na jiných hodnotách v záznamu. Nejjednodušší způsob spočívá v nahrazení průměrnou hodnotou podle subpopulace, kam subjekt patří. Výhodou imputace pomocí regresní analýzy je, že dává správnější odhady. Tento způsob je však výpočetně náročný a vede k tomu, že proměnné po imputaci jsou „lépe“ korelované než v původnosti.

2.6 Statistika a modelování

Statistiku můžeme definovat i jako vědu o variabilitě empirických dat. Variabilita dat má různou příčinu. Především se zajímáme o systematické vlivy a efekty faktorií, které sledujeme, abychom jim lépe porozuměli a mohli je případně

PŘEHLED STATISTICKÝCH METOD

ovlivňovat. Tato složka variability je však překryta jinou variabilitou, jež má příčinu v nepřesnosti měření, v intraindividuální variabilitě a v dalších náhodných vlivech. K tomu se ještě přidává výběrová chyba vznikající tím, že data pouze zastupují obvykle větší univerzum případů a situací. Studie mají odhalit, potvrdit či vyvrátit obecnější zákonitosti, které přesahují aktuálně dostupná data. Data – resp. výběrový soubor – nesouhlasí přesně s celou populací v některých charakteristikách, jichž se mají naše poznatky týkat. Zde je důležité rozlišovat mezi vysvětlenou a nevysvětlenou variabilitou. Obvykle hledáme nějaký význam v datových konfiguracích. Identifikujeme základní vzorce variability. Nevysvětlená variabilita je to, co zůstává, když odečteme vše, co považujeme za významné vlivy. Nevysvětlenou variabilitu modelujeme pomocí nějakého náhodného procesu. Užitečnost takového přístupu mnohdy nezávisí na tom, zda jsou naše pozorování skutečně náhodná. Jedná se o naši odpověď na komplexitu dat.

Úlohu statistiky ve výzkumu nám přibližuje představa o **signálu a šumu**. Představujeme si, že data mají tvar

$$\text{data} = \text{signál} + \text{šum},$$

kde signál představuje informaci o objektech, jež zkoumáme. Toto schéma využívá, že signál, který vyjadřuje určité pravidelnosti, je více či méně zakryt šumem.

Data mohou být interpretována, pokud máme k dispozici jejich model, jehož hlavní část tvoří nějaká rovnice. Pomocí modelu se snažíme vystihnout podobu signálu. Data ve statistice modelujeme podle schématu

$$\text{data} = \text{predikce modelem (funkce)} + \text{reziuální hodnota}.$$

Model zjednodušeně zrcadlí strukturu dat. Reziuální hodnota nebo chyba reprezentuje diferenční mezi modelem a napozorovanými daty. Je důležité, aby reziuální hodnota neobsahovala žádnou evidentní strukturu. Pokud ji obsahuje, je nutné model dále upravovat, až model zahrne všechny identifikovatelné struktury signálu.

Pravděpodobnostní modelování dat při analýze pomáhá vytvořit model dat, popsat povahu signálu, charakterizovat šum a usuzovat pomocí modelu o mechanismech, jež data generují. Modelování je důležitou činností při využívání statistických a pravděpodobnostních metod. Vyjádříme přesněji, o co jde. V modelování usilujeme o zjednodušenou reprezentaci reality. Modelem je např. rovnice popisující volný pád tělesa $s = 0.5at^2$ ve vakuu, kde a je gravitační konstanta. Jde o deterministický model. V pravděpodobnostních modelech uvažujeme vždy nějakou náhodně proměnlivou složku modelu. Zvláště jednoduchým modelem

Obr. 2.7 Proces statistického modelování



ve statistice je jednovrcholový symetrický tvar křivky rozdelení dat neboli Gaussova křivka. Pro jednoduchý popis takových dat stačí vypočítat jejich průměr a směrodatnou odchylku jako míru rozptýlenosti.

Obvykle rozlišujeme při modelování fáze základní formulace modelu, přizpůsobení modelu datům (parametrizace modelu), jeho kritiku, resp. vyhodnocení a interpretaci. Tyto fáze a jejich návaznosti ukazuje obrázek 2.7.

V první fázi popisujeme proces, který nás zajímá. Specifikujeme pravděpodobnostní model, jenž bude využívat určité předpoklady o chybách měření, o vztazích mezi proměnnými, o způsobu sběru dat atd. Tyto volby budou záviset na našich teoretických poznatkách a zkušenostech. Druhá fáze znamená, že některé vlastnosti modelu budeme upřesňovat pomocí pozorovaných dat. Provedeme přitom kvantifikaci vztahů a náhodné složky. Odhadujeme parametry modelu. Kritika modelu spočívá ve vyhodnocení adekvátnosti modelu. Ptáme se, zda se jedná o přesnou a úspornou reprezentaci modelovaného fenoménu. Také se ptáme, zda předpoklady modelu jsou splněny. Přitom můžeme srovnávat nový model s jinými modely a používáme různé diagnostické metody. Při revizi modelu model doplňujeme o další prvky nebo ho naopak zjednodušíme (využíváme princip Occamovy břitvy). Poslední fáze tvoří využití modelu při řešení výzkumných otázek.

Při modelování si připomínáme slova statistika G. E. P. Boxe: „Všechny modely jsou špatné, ale některé jsou užitečné.“

Souhrn

Statistika je naukou o získávání kvantitativních údajů (dat) a o jejich analýze. Statistická analýza je prostředkem k získání závěrů o údajích, jež se získaly v rámci empirického výzkumu. Každá statistická studie zaznamenává údaje o sledovaných jednotkách (jedincích, věcech, zvířatech, institucích nebo specifikovaných jevech) tak, že u nich zjišťuje hodnoty proměnných. Nejdůležitějším aspektem statistické studie je způsob, jak se data získala. V výběrových šetřeních získáváme data bez narušení procesu, v němž vznikají. Jednorázové průřezové šetření je důležitým typem výběrového šetření. Vycházíme přitom ze vzorku specifické populace, abychom získali informaci o celé populaci. V experimentu uplatňujeme ošetření jednotky a měříme jeho efekt. Snažíme se zjistit, zda zásah (intervence) vyvolá předpokládaný účinek. Na obrázku 2.8 jsou zachyceny základní prvky výzkumu pomocí statistiky a vztahy mezi nimi.

Konceptuální mapa (podle Lajoie, 1998, s. 95) uvádí do vztahu čtyři prvky statistického zkoumání (položení statistické otázky, sběr dat, analýza dat a interpretace) a další statistické pojmy. Většinu z nich jsme zatím jen naznačili. Probíráme je v dalších kapitolách. Statistické otázky se týkají popisu dat, zkoumaní rozdílů mezi množinami dat, zkoumaní závislosti mezi proměnnými a zobecňování závěrů na populaci. Problém sběru dat obsahuje určení zkoumaných populací a metod sběru dat. Analýza dat se týká popisu dat, komparace dat, hledání vztahů. K dispozici jsou metody grafické, numerické, použití tabulek a metody statistického usuzování (statistické testy a intervalové odhady). Interpretace výsledků nás vede zpět k položeným otázkám. Ptáme se, jak nám data pomáhají zodpovědět jednotlivé otázky.

Plánování výběrového šetření a experimentu se věnuje podrobně Bortz (1995) a Ferjenčík (2000). Problematiku modelů a modelování na obecné úrovni skvělým způsobem postihuje antologie textů, kterou připravili Berka a Tondl (1967).

Obr. 2.8 Konceptuální mapa výzkumu pomocí statistických metod

