

Tab. 2.2 Účinek rušivého faktoru – podrobnější data o nehodách

	Nízká rychlost			Vysoká rychlost		
	Ochranný pás		Součet	Ochranný pás		Součet
	použitý	nepoužitý		použitý	nepoužitý	
Řidič zemřel	4	2	6	6	18	24
Řidič nezemřel	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 ukazují relativní četnosti 40/60 a 10/40. Rychlost nárazu je v tomto případě rušivým faktorem, protože maskuje efekt – pokud je nějaký – faktoru „použití ochranného pásu“. Vliv rušivého faktoru vede ke zkrácení 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 mačování – podrobněji kap. 2.4.3).
- Omezení působení neznámých potenciale matoucích proměnných docílíme 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).
- Znamé 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 měřítko. Navíc vyhovuje požadavku, že vzdálenosti jednotlivých údajů uvnitř tohoto systému jsou dány nějakou jednotkou měření. Proto je možné je počítat a odčítat.

Měřítka z největší informační hodnotou je **poměrové měřítko**, které má všechny vlastnosti intervalového měřítko 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:

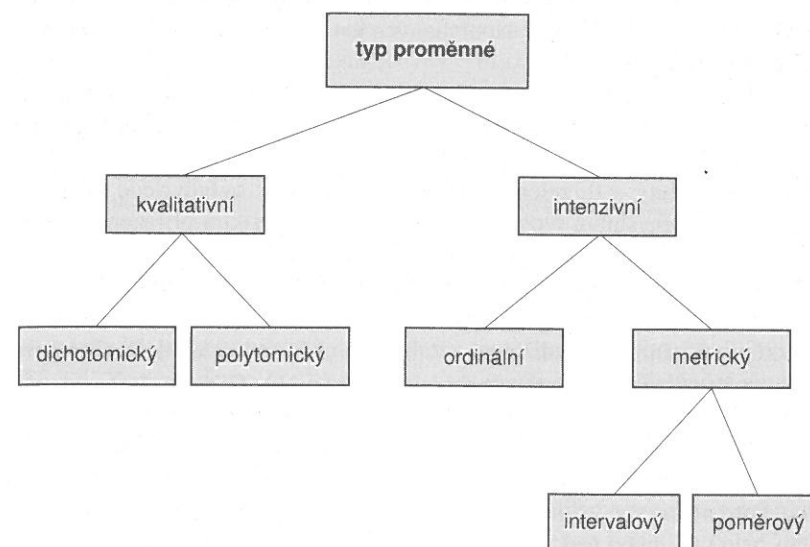
- Při *nominálním měřítku* můžeme mít pouze vztahy $x_a = x_b$ nebo $x_a \neq x_b$ (např. muž, žena).
- *Ordinální měřítko* také dovoluje vztahy $x_a > x_b$ nebo $x_a < x_b$ (např. světlý, šedý, tmavý).
- *Intervalové měřítko* navíc předpokládá, že může být definována velikost rozdílu, takže objekt A se liší o $x_A - x_B$ jednotek od objektu B (např. teplota ve stupních Celsia).
- *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 nehraje 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 spojitá proměnné

Při zpracování dat hraje roli 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
tepová frekvence	metrický spojitý i diskretní
počet vstřelených gólů	metrický diskretní
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 diskretními údaji pracovat jako se spojitými, např. tepovou frekvencí můžeme považovat za spojitou proměnnou, pokud se setkáváme v analýze s velkými rozdíly tepových frekvencí.

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

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

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

2.3 Kvalita měření

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

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

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

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

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

$$Y = T + E,$$

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

2.3.1 Objektivita

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