STATISTIKA – LETMÉ NAHLÉDNUTÍ DO VÝKLADOVÉHO SLOVNÍKU

prof. RNDr. Stanislav Komenda, DrSc.
Oddělení biometrie LF UP Olomouc

Statistika doprovází lidskou civilizaci od samého jejího počátku. Žádný státní nebo městský útvar se neobešel bez toho, že by měli ti, kdo ho řídili, *informaci* o poměrech v řízeném systému (počtu obyvatel, jejich majetku, včetně domů a dobytka) a mohli tak hodnotit své možnosti (zahájit loupeživou válku). Předmětem zájmu statistiky proto není informace o jednotlivostech a jednotlivcích, ale informace o celcích, souborech, populacích.

Pořizování a vedení statistik je podmíněno možností *měření.* Před asi třemi sty lety se objevují pokusy měřit zdravotní stav obyvatelstva (londýnské statistiky) a postupně tak vzniká *zdravotnická statistika.* Zatímco lékař u svého klienta konstatuje, zdali je nemoc přítomna anebo ne, měří zdravotnická statistika *morbiditu, nemocnost* na škále od 0 do 100 %.

Až do konce 19. století se vycházelo z představy, že to, co statistika uvádí, popisuje stav celé příslušné *populace.* Teprve před sto lety vzniká *teorie výběru*; matematicky bylo dokázáno, že – za jistých předpokladů – je schopen i poměrně omezený vzorek *náhodně* vzatý z *referenční populace* vypovídat o stavu, tj. ukazatelích v této populaci s dostačující *přesností.* Začala se vyvíjet *matematická (inferenční, induktivní) statistika,* zabývající se studiem *matematických modelů usuzování* ze vzorku na celek. Jejími hlavními formami jsou *statistické odhadování parametrů pravděpodobnostních rozdělení* a *testy statistické signifikance.*

Nová představa se nijak nedotkla významu měření; naopak, prohloubila jej. Měřicí škály statistických *ukazatelů* se třídí od jednoduchých a informačně chudších k složitějším a informačně bohatším. *Měření nominální* umisťuje měřené objekty či subjekty do množiny *úrovní* bez vnitřní uspořádanosti (*znak* Pohlaví jedince – s úrovněmi Muž, Žena; *znak* Krevní skupina – s úrovněmi 0, A, B, AB). *Měření ordinální* přiřazuje měřeným objektům úrovně vybírané z množiny, jejíž prvky jsou uspořádány vzestupně či sestupně (*znak* Položka dotazníku s typickou např. pětibodovou stupnicí – naprosto nesouhlasím, v podstatě nesouhlasím, je mi to jedno, v podstatě souhlasím, naprosto souhlasím). Obojí uvedená měření se označují jako *měření kategoriální.* Informačně bohatší jsou *měření metrická,* kdy jsou stavům sledovaného ukazatele přiřazována reálná čísla. Metrické měření má svou *fyzikální jednotku.* Podle povahy věci se rozlišují *měření intervalová,* kdy není jednoznačně definována *nula* používané stupnice (teplota měřená ve stupních Celsia nebo ve stupních Fahrenheita) a *měření poměrová,* kdy nula jednoznačně definována je (somatická, fyziologická, biochemická a jiná klinická měření). Ukazatele (znaky), které jsou předmětem sledování, bývají podle definice povahy *spojité* či *nespojité (diskrétní).* Stačí představit si znak „počet dětí v rodině“ oproti znaku „tělesná výška jedince“. Znak ex definitione spojitý se ovšem prakticky vždy měří nespojitě, s jistou zvolenou *rozlišovací schopností.* Volba je přitom záležitostí účelovou; tělesnou hmotnost měříme v kilogramech, detailnější měření by byla práce s *náhodnými čísly.* Tato volba má jenom málo společného s *přesností měření.*

Výsledky měření a pozorování tvoří pak *data.* První operací s daty je *třídění;* ještě před patnácti lety nejhrubší a časově nejvíce náročnou práci obstarává dnes výhradně *počítač.* Data jsou vytříděna do podoby *datové matice, databáze,* obvykle za pomoci *tabulkového procesoru* EXCEL, instalovaného dnes běžně i v klinických počítačích. Obvyklý způsob záznamu vymezuje řádky pro pacienty a sloupce pro znaky (ukazatele). S databází instalovanou na disketě nebo CD se případně kontaktuje středisko *biostatistického (biometrického) zpracování dat.*

Standardní součástí statistického hodnocení dat je výpočet *popisné (deskriptivní) statistiky.* Ten navazuje bezprostředně na třídění dat. V případě dat kategoriální povahy je výsledkem statistického popisu *četnostní tabulka,* obvykle se dvěma vstupy – se dvěma znaky (úrovně jednoho vstupují jako řádky, úrovně druhého jako sloupce), jejichž případná *souvislost* nás zajímá. Políčka tabulky uvádějí *absolutní četnosti výskytu* objektů nebo subjektů s úrovněmi znaků odpovídajícími danému řádku a danému sloupci. Hlavním důvodem provádění statistického popisu dat je jejich *redukce a tedy koncentrace informace* v datech obsažené. Četnostní (frekvenční) tabulka obsahuje zpravidla několikanásobně méně údajů ve srovnání s objemem původních dat. To má zásadní význam pro manipulaci s informací, pro její předávání a sdělování. Redukce dat v případě údajů metrických se děje výpočtem *základních statistických charakteristik* souboru hodnot daného ukazatele, tj. hodnot příslušného sloupce v databázi.

Statistické charakteristiky polohy souboru dat vypovídají o hodnotě, kolem níž se data soustřeďují. Výpočet respektuje povahu měřicí škály. Pro měření kategoriální se užívá *modu a mediánu,* pro měření metrická především *aritmetického průměru.*

Statistické charakteristiky variability souboru vypovídají o míře rozkolísanosti, proměnlivosti dat kolem středu souboru. Jednoduché charakteristiky vycházejí z *minimální a maximální hodnoty v souboru;* pro metrická měření je nejpoužívanější charakteristikou variability *rozptyl* a z něho odvozená *směrodatná odchylka,* případně *variační koeficient.* Ustálilo se uvádění redukovaných charakteristik souboru metrických dat tím, že jsou místo primárních dat uváděny pouhé tři údaje – *rozsah souboru, průměr a směrodatná odchylka.*

Dalším typem statistických charakteristik souboru dat jsou *charakteristiky souvislosti dvou ukazatelů.* V případě kategoriálních dat vychází jejich výpočet z četnostní tabulky. Velmi často se přitom využívá statistiky *chí-kvadrát. Míry souvislosti, závislosti či asociace měří stupeň této vazby mezi*

ukazateli na stupnici od nuly do jedné. V případě dat metrických se v této roli používá *korelační koeficient;* bylo jich navrženo několik. Detailněji studovat závislost jedné *veličiny na veličině jiné* můžeme pomocí *regresní analýzy. Statistický výpočetní software* nabízí celou řadu metod použitelných pro biostatistické hodnocení dat.

K základním konceptům využívaným v matematické statistice patří pojem *pravděpodobnosti.* Na rozdíl od statistiky je *teorie pravděpodobnosti* matematická disciplína poměrně mladá – tento koncept byl vymezován od 16. století. Ačkoli helénská filozofie formulovala snad všechny otázky dnešní filozofie, s konceptem *nahodilosti* si neporadila. Pro biometrickou aplikaci je důležité vědět, že pravděpodobnost je *mírou možnosti výskytu jevu;* musí jít o tzv. *náhodný jev,* což souvisí s konceptem *hromadnosti, masovosti.* Pravděpodobnostní představy nelze aplikovat na *jevy jedinečné.* Dále je podstatné vědět, že pravděpodobnost výskytu jevu jako teoretický koncept lze specifikovat (vlastně odhadovat) číselně pomocí *relativní četnosti,* s níž se tento jev vyskytuje v souboru realizací (*opakování)* příslušného pokusu, tj. souboru opakovaných pozorování, při každém z nichž se sledovaný jev může nebo nemusí vyskytnout. Základním požadavkem je přitom *nezávislost získaných dat.* Tak např. *morbidita* je odhad pravděpodobnosti výskytu dané nemoci v referenční populaci.

Hlavní *heuristickou metodou* při získávání nových poznatků je *srovnávání, komparace.* Účinek zákroku, nového léku nebo ošetřovacího postupu se neposuzuje podle nominálních hodnot příslušného ukazatele, ale z výsledku srovnání souboru hodnot tohoto ukazatele naměřených po aplikaci zákroku se souborem hodnot téhož ukazatele naměřených v situaci *kontrolní,* která je shodná se situací *pokusnou* ve všech ohledech s výjimkou onoho zákroku. Aby se dosáhlo co nejvěrnější shody okolností, využívá se *placebo.* Aby subjekt nerozpoznal, zda byl zařazen do skupiny pokusné nebo do skupiny kontrolní, bývá pokus *zaslepen* případně *dvojitě zaslepen* (nejen subjekt, ale ani pracovník hodnotící případný účinek zákroku by neměl vědět, do které skupiny patří subjekt právě hodnocený).

Existuje komplex pravidel pomáhajících uskutečňovat *kontrolu* i ve více či méně složitých situacích. Tato pravidla formulují *plány pokusu.* Je třeba rozlišovat mezi plány *prospektivními a retrospektivními, studiemi longitudinálními a průřezovými (transverzálními).* Tyto plány usilují zachytit vliv času a tedy opakování měření, zejména pak oddiferencovat variabilitu *interskupinovou a intraskupinovou.*

Odhadování parametrů teoretických modelů pravděpodobnostních rozdělení mívá formu *odhadů bodových a odhadů intervalových. Interval spolehlivosti* usiluje na základě dat a předpokladu, že se studovaný ukazatel řídí určitým zákonem rozdělení, číselně specifikovat toto rozdělení tak, aby optimálně odpovídalo empirickým pozorováním; specifikuje interval, který by měl s dostatečně vysokou pravděpodobností *pokrýt* neznámou hodnotu odhadovaného parametru. Pomocí tohoto postupu lze konkretizovat koncept *přesnosti měření;* přesnost má dvě složky: *spolehlivost* (danou předem, obvykle pravděpodobností 0,95) a *ostrost* (specifikovanou šířkou intervalu – užší interval znamená vyšší přesnost).

Testy statistické významnosti jsou nejčastěji aplikovaným postupem *statistické indukce.* Problém, který má být pokusem řešen, se formuluje jako *nulová hypotéza* (studovaný zákrok nemá na statistické chování ukazatele v referenční populaci vliv). Za předpokladu její platnosti se matematicky odvodí, jak by se mělo chovat *testové kritérium*, definované tak, aby se v jeho chování obrážel případný vliv zákroku. Chová-li se testové kritérium podle *očekávání,* hypotéza se podrží – vliv zákroku na sledovaný ukazatel prokázán nebyl. Je-li statistické chování testového kritéria ve významném rozporu s očekáváním, *nulová hypotéza se zamítá jako nesprávná.* Výsledek testu se prohlásí za *statisticky signifikantní.* Tato procedura statistické indukce je vystavena *riziku omylu. Chyba 1. druhu (pravděpodobnost jejího výskytu, signifikance,* bývá obvykle specifikována hodnotou 0,05) představuje jev, kdy je nulová hypotéza zamítnuta jako nesprávná, ačkoli ve skutečnosti platí. *Chyba 2. druhu* představuje jev, kdy je jako správná neoprávněně přijata nulová hypotéza ve skutečnosti neplatná; pravděpodobnost jejího výskytu se běžně v software nepočítá. Jde o velmi sofistikovaný koncept, protože tvrzení vyslovené nulovou hypotézou se může od neznámé skutečnosti odlišovat celou škálou diferencí. Úplný popis rizika testu pak představuje celá křivka pravděpodobností jako funkce těchto diferencí; nazývá se *silofunkce statistického testu.*

Běžně se v biometrii používá několik desítek testů. Pokusy se složitější strukturou srovnávání se vyhodnocují *analýzou rozptylu (ANOVA)*, s rozsáhlým počtem variant. Jistou modifikací je známý *t-test podle Studenta.*

Rozsáhlá třída testů statistické významnosti reaguje na poměrně přísné formální požadavky na typ pravděpodobnostního rozdělení sledovaného znaku tím, že primární data nahrazuje jejich *pořadími (ranky). Neparametrické (pořadové) testy* jsou obecněji aplikovatelné než *testy parametrické.*

Ne všechny testy jsou založeny na principu srovnávání. Důležité a často používané jsou i testy týkající se závislosti veličin, testy hypotéz předpokládajících, že *korelace veličin v referenční populaci je rovna nule.*

Statistické multivariační metody řeší řadu tzv. *heuristických problémů* formulovaných ve vícerozměrném prostoru sledovaných veličin. Jejich aplikace je bez použití speciálního výpočetního software prakticky vyloučena.