

Zkoumání životního způsobu se týká i třetí a pravděpodobně nejzajímavější z této série teoretických příspěvků — *K využití kategorie času při zobrazení životního způsobu* autorů G. Lippolda a G. Manze. Po vstupní diskusi pojmů životního způsobu a životního stylu (životní styl je jedním z projevů způsobu života) přecházejí autoři k teoreticky i metodologicky podnětnému rozboru vztahů mezi kategoriemi času a socialistického životního způsobu, resp. mezi těmito kategoriemi a časovým rozpočtem obyvatelstva. Společenské aktivity, jimiž se způsob života projevuje, se vždy odehrávají v čase a prostoru. Nejlepší metodou, jak životní způsob zachytit, je postihnout časový průběh lidských aktivit. Autoři proponují hierarchii kategorií, které je třeba při tomto sledování uplatnit, a kriticky rozebírají možnosti a meze techniky výzkumu časových rozpočtů obyvatelstva.

Uvedený příspěvek představuje přechod k dalšímu tematickému bloku statí s metodologickou problematikou. H. Berger a W. Klimek se ve statí *K pojmu a úloze indikátorů v sociologickém výzkumu* zabývají pojmem indikátoru a jeho gnoseologickou funkcí v sociologii, podávají přehled o využití sociálních indikátorů v řízení socialistické společnosti a kriticky hodnotí „hnutí sociálních indikátorů“ v západních zemích. Indikátory v sociologickém výzkumu definují jako „teoreticky vymezené a empiricky ověřené, resp. ověřitelné ukazatele pozorovatelných sociálních jevů a procesů. Indikátory jsou objektivními koreláty sociologických pojmů“. Důležitost problematiky sociálních indikátorů plyne z jejich významu pro sociální politiku a pro řešení gnoseologických aspektů sociologického výzkumu. Autoři rozlišují — podle našeho názoru poněkud problematicky — mezi sociologickými a sociálními indikátory (prvé se vztahují k „sociologickým“ výzkumným objektům, druhé k sociálním jevům, které nejsou zařazeny do sociologického výzkumného procesu) a mezi indikátory a ukazateli (za ukazatele zřejmě považují prvky pojmového modelu zobrazovaného objektu). Východiskem pro konstrukci soustavy sociálních ukazatelů na celospolečenské úrovni by se podle nich měl stát pojem životního způsobu. V této souvislosti lze vzpomenout na československé diskuse, které naznačily, že samotná kategorie životního způsobu zřejmě pro vymezení procesů sociálního rozvoje nepostačuje.

Na konkrétní úrovni ilustruje postup při vytváření indikátorů stať E. Prillera *Konstrukce ukazatelů pro sociologický výzkum pracovních podmínek*. Autor vychází z pojmové analýzy kategorie pracovních podmínek a deduktivní cestou se postupně dopracovává ke strukturnímu schématu (operacionálnímu modelu), jež by se mělo stát bezprostředním východiskem pro tvorbu indikátorů. Vlastní úkoly tvorby ukazatelů se v článku neřeší.

Stručně, avšak velmi přehledně a informativně je koncipován příspěvek W. Hartmanna a R. Welskopf *K využití vícerozměrného škálování v sociologii*. Učebnicovou formou shrnuje funkce a principy vícerozměrného škálování a ilustruje je na příkladu z výzkumu pracovní motivace.

Jako technologický proces analyzuje postup sociologického výzkumu F. H. Schröder v článku *Problémy vytvoření účinné výzkumné technologie*. Inspirován technickými a systemologickými přístupy navrhuje řadu schémat „technologické struktury“ výzkumného procesu a uvažuje o využití techniky síťového grafu při jejím projektování. Podobně jako v minulosti u mnoha jiných případů technického analogizování se klade i zde otázka, jaký poznávací zisk vlastně vyplynul z přenášení technické terminologie do oblasti sociologického bádání.

Poslední blok původních příspěvků je věnován kritice buržoazních sociologických koncepcí. H. Steiner v jedné z nejlépe napsaných kapitol celé ročenky, *Sociálně politická orientace — východisko z krize buržoazní sociologie?*, vychází z předpokladu, že buržoazní sociologie je již od svého vzniku plodem krize kapitalistické společnosti. Nedávny návrat této sociologie k problémům sociální politiky je výrazem snahy lépe plnit ideologické a sociotechnické funkce v období prohlubující se společenské krize. Článek je výborně dokumentován aktuální západoněmeckou literaturou. Úlohou životního způsobu v ideologické konfrontaci socialismu a kapitalismu se zabývá H. Lehmann. V článku *Způsob života v soudobém ideologickém třídním boji ze sociologického hlediska ukazuje*, jak jsou kategorie životní způsob a „kvalita života“ využívány v propagandě proti socialistickým zemím.

Řada zajímavých informací je obsažena i ve zbývajících částech ročenky, přinášejících jednak zprávy o činnosti vědeckých rad pro sociální politiku a demografii, pro sociologický výzkum a pro ekonomický výzkum, jednak recenze domácích, sovětských a západoněmeckých prací. Ročenka zde plní úlohu časopisu, který sociologové v NDR dosud k dispozici nemají.

Publikaci *Ročenky pro sociologii a sociální politiku 1980*, a zejména záměr vydávat ročenku pravidelně i v dalších letech, je třeba ocenit jako záslužný a zdařilý počín. Sociologie u našich severních sousedů tak získává reprezentační publikační příležitost, kterou u nás již dlouhá léta postrádáme a kterou časopis plně nahradit nemůže. Na vrub skutečnosti, že jde o první vydání ročenky, lze zřejmě přičíst i určitou všeobecnost, rámcový charakter značné části zveřejněných příspěvků a absenci konkrétnější sociologické tematiky, zejména ve vazbě na provedené výzkumy. V ročence také zatím není výrazněji zastoupena problematika sociální politiky, kterou avizuje její název. K těmto otázkám se však redaktori jistě dostanou v dalších vydáních ročenky, která budou se zájmem očekávat i českoslovenští čtenáři.

Michal Illner

Diskriminační analýza – klasický nebo diskrétní model?

M. Goldstein – W. R. Dillon: *Discrete Discriminant Analysis* (Diskrétní diskriminační analýza)

New York, John Wiley & Sons, 1978, X + 186 s.

Diskriminační analýza je jako statistická analytická metoda často používána pro analýzu (většinou komparační) sociologických dat. Vzhledem k úspěšnosti aplikací, jistě atraktivnosti modelu, snadné programovatelnosti pro počítače jde o metodu stále oblíbenější. Avšak čím častěji aplikujeme nějakou metodu, tím spíše objevíme nejrůznější nejasnosti, otázky, nepříjemné a nepochopitelné aspekty. To je pak signálem k dalšímu metodologickému bádání, k modifikacím základního modelu, ke konstrukci modelů nových. Každý, kdo se trochu více a hlouběji zabýval aplikacemi metody diskriminační analýzy ve společenských vědách, si všiml dvou nepříjemných aspektů: a) malé stability výstupních parametrů (koeficientů diskriminačních funkcí), b) nepochopitelných hodnot znaménka koeficientu diskriminačních funkcí. Tyto dva nepříjemné aspekty se zdají být průvodním zjevem při použití diskrétních (i dichotomických) dat. A odtud také začíná kniha Goldsteina a Dillona, odtud je argumentována potřeba vyvíjet samostatné modely diskriminační analýzy pro nominální, ordinální ale i diskrétní kardinální data.

Kniha je stručná, s jednoznačným zaměřením na statistickou praxi. Aplikační argumentace je vzata většinou ze sociálních věd. Práci lze charakterizovat třemi hlavními rysy, které mají pro čtenáře-sociologa různý význam:

1. Je použito formálního jazyka matematiky a jednotlivé vlastnosti modelů nejsou nijak podrobně verbálně vysvětleny. Autoři předpokládají, že čtenář má profesionální matematické znalosti. Nematematický čtenář však může získat přehled o metodách z četných příkladů, které jednotlivé postupy dokumentují. *Matematický a formální* aspekt není předmětem této recenze. Poznamenejme jen, že uvedení všech metod a modelů je jasné a přehledné. Z tohoto hlediska lze charakterizovat knihu jako úspěšnou přehledovou práci, shrnující závažný dílčí problém.

2. Celá třetina knihy (58 stran) je věnována přetisku *počítačových programů*. To by v oblasti sociologie zřejmě předpokládalo anotaci s uvedením obsahu programů a jejich technických parametrů. Je tu zařazeno šest samostatných programů, napsaných v jazyce FORTRAN IV jako soubory subrutin (jsou též závislé na subrutinách knihovny SSP počítačů IBM). Pět programů má uspišit a zjednodušit aplikaci metody popsané v knize.

3. Vlastním důvodem pro recenzi knihy je její *aspekt metodologický*. Výsledky uváděné v první kapitole a poté implicitně zahrnuté ve všech dalších kapitolách mají metodologické důsledky pro aplikaci diskrétní diskriminační analýzy i pro aplikaci klasického lineárního modelu vůbec. Kniha se tak stává odrazovým můstkem pro metodologické úvahy, které mohou bezprostředně ovlivnit rozvoj metodologie i metodiky analýzy sociologických dat. Tato recenze je proto především pokusem o uvedení závěrů, které autoři zahrnuli do své knihy v jazyce sociologickému uživateli nesrozumitelném, ale které jsou v knize jednoznačně modelově *dokázány* nebo simulacemi a praktickým rozbořem dříve publikovaných analýz prokázány.

Hovoříme-li o diskriminační analýze, máme na mysli téměř vždy klasický lineární

model založený na lineárních diskriminačních funkcích a na předpokladu mnoho-rozměrného normálního rozložení spojitých náhodných veličin. Tento model je dobře znám, matematicky je velmi podrobně prozkoumán, je prověřen mnoha simulačními experimenty i aplikační praxí. Přesto (jak už bylo uvedeno) se při jeho použití setkáváme s některými problémy, které u nezkušeného uživatele mohou vést k chybným závěrům a mohou zavést i uživatele zkušeného, neboť neadekvátnost výstupu metody může být značně zamaskována.

Původní model diskriminační analýzy slouží k rozlišování statistických jednotek podle předem známých tříd (typů, diagnóz, skupin), a to na základě nepřímých indikátorů. Ve společenských vědách je takový rozhodovací aplikační postup relativně řídký. Častěji je používán v komparativních postupech (vytváření modelu rozlišovacích indikátorů mezi jednotlivými předem danými třídami), resp. i pro účely vytváření typů. Tedy původní prakticko-diagnostické uplatnění metody se mění na uplatnění explorativně teoretické. To ovšem s sebou přináší nové úlohy pro metodologii i pro matematické ověření. Komparační aplikace jsou založeny na interpretaci koeficientů diskriminačních funkcí (na jejich absolutní hodnotě i znaménku). Metoda je založena na tom, že hledá diskriminační funkce tak, aby určitá (předem zvolená) kritéria byla optimální. V blízkosti optimální hodnoty však může existovat řada dalších řešení, která vedou na jiné diskriminační funkce (jiné jejich koeficienty, a tedy i na jiné komparační interpretace), přičemž kvalita rozlišení, která je kritériem, je téměř stejná. Je proto možné, že výstup z metody je závislý na empirických nahodilostech výběrového souboru, které vedou k upřednostnění jednoho z mnoha podobných řešení, a skrývá nebezpečí interpretačních artefaktů. Toto nebezpečí se týká též metod výběru proměnných diskriminačního modelu. Statistická kritéria zde nemohou pomoci k rozhodnutí, zda řešení přijmout, nebo ne. To umožňuje jen kontrola jinou statistickou metodou, empirický cit uživatele a především teoretické úvahy zasazující řešení jednotlivého případu mezi ostatní zobecněné poznatky vědy a podobné empirické analýzy jiných šetření. Uvedená zásadní potíž aplikace modelu plyne z toho, že při empirické klasifikaci statistických jednotek (diagnostika) lze s úspěchem používat zprostředkovaných či nepravých korelací pro určení empirické indikace skupiny (typu diagnózy), pro teoretickou zobecňující interpretaci je však nutné znát skutečné, čisté korelace, aby bylo možné odhalit pravý význam vah v diskriminační funkci a tím i skutečný význam jednotlivých indikátorů.

Vraťme se však ke knize. Při popisu jednotlivých kapitol uvedeme i metodologické aspekty, jimiž se M. Goldstein a W. R. Dillon zabývají: jsou to především vlastnosti vycházející z porušení předpokladu spojitosti (a tím i normality a linearity) vstupních proměnných, neboť ty se přímo váží k tématu knihy. Praktické problémy, s nimiž se tak často setkáváme při aplikaci diskrétních dat v diskriminační analýze, jsou shrnuty v textu, příkladech a modelech první kapitoly — *Lineární diskriminační funkce*. Běžné rutinní postupy jsou založeny na Fisherově přístupu, který je také v první části kapitoly popsán. Předpoklady správnosti postupu: průměry, rozptyly a lineární korelační koeficienty dostatečně dobře charakterizují rozložení dat a vztahy mezi daty ve všech jednotlivých skupinách, které chceme rozlišovat. (V knize se autoři zabývají vlastně rozlišováním pouze dvou různých skupin, tedy rozebírají matematicky i metodologicky nejjednodušší případ.)

Ve třetí části kapitoly (*Užití kvalitativních dat*) uvádějí autoři zcela konkrétní model pro binární data (vzatá z výzkumné aplikační praxe), v němž nelze rozlišit za pomoci klasické diskriminační analýzy dva velmi důležité typy mající nemonotónní definici (0,1) a (1,0) proti (1,1) a (0,0), tedy nekonzistentní vs. konzistentní případ pro obě dichotomické indikace). Už tento model nutí uživatele k vážnému zamyšlení nad aplikacemi metody, neboť právě takovéto případy jsou v aplikacích časté. Metodu lze zřejmě používat pro binární data jen tam, kde indikační proměnné vytvářejí rozhodnutí souběžně.

V první kapitole jsou dále rozebrány jiné dva důležité aspekty aplikace diskrétních dat:

a) Nominální diskrétní data musí být před aplikací dichotomizována na tzv. pomocné („dummy“) proměnné, a tudíž v nich nastává problém inkonzistence typů uvedený výše. U ordinálních proměnných lze provést pomocné pseudokvantifikace (nejčastěji číslo pořadí kategorie znaku), což však může buď do dat vnést umělé linearity, nebo skutečné linearity zakrýt. Oproti modelu nezávislosti nebo monotonní závislosti je u diskrétních dat interakční efekt daleko častější (a mívá přirozený vznik a vysvětlení, nelze jej proto bez následků z analýzy vypustit). Tyto vlastnosti diskrétních dat vedou ke zvláštnímu jevu, který pozoroval každý, kdo s metodou častěji pracoval: koeficienty diskriminační funkce mají opačná znaménka, než jak vyžaduje logika věci a ukazují jiné analýzy.

b) Aby koeficienty byly interpretačně užitečné, musí být stabilní, to jest invariantní vzhledem k různým pravděpodobnostním výběrům. Autoři uvádějí výsledky společenskovední studie, v níž experimentovali s tím, že kromě úplného náhodného výběru použili pro základ interpretace pouze 70 % tohoto výběru; výsledky se lišily velmi značně. Podle jejich interpretace tu je nízká stabilita způsobena diskrétními hodnotami a užitím kódů kategorií ordinálního znaku jakožto kvantifikace.

Rozbor z první kapitoly je dostatečně přesvědčivý a argumentativní pro vývoj speciálních metod založených na vlastnostech diskrétních dat. Zároveň je také imperativem k pozornosti při používání metody. Celá další část knihy se pak zabývá shrnutím současného stavu diskrétní diskriminační analýzy pro rozlišení dvou tříd. Především jsou diskutovány dichotomické indikátory.

Kapitola 2, *Diskrétní klasifikační modely*, uvádí obecný model přiřazování statistických jednotek na základě diskrétních indikátorových proměnných. Nejprve je zaveden úplný multinomický model (to jest čistý případ podle základní teorie). Tento model je však nereálný pro větší počet proměnných vzhledem k velkému počtu polí mnohorozměrné tabulky, která vzniká kombinací všech indikátorů a nízkých i nulových četností v nich. Důsledkem je nestabilní klasifikace, protože odhady pravděpodobností pro jednotlivá pole jsou velmi nepřesné. Největší potíž činí nulová pole, která jsou nulová empiricky, nikoli teoreticky – nejde o tzv. strukturní nuly. Tento základní model zároveň odkrývá slabiny diskrétního přístupu: nelze používat velký počet diskriminujících indikačních proměnných a jsou vyžadovány velké výběrové rozsahy v jednotlivých skupinách.(1)

Úplný multinomický model je pro praxi naprosto nevhodný a proto hledáme jeho modifikace a příbuzné přístupy tak, abychom buď snížili počet odhadovaných parametrů, nebo využili pro odhad četností některých strukturních vlastností úplného rozložení všech znaků.

Kapitola dále uvádí celou řadu specifických a velmi užitečných modelů, které podstatně zlepšují postup: A) modelové reprezentace pro dichotomická data – Bahadurův model, logaritmicko-lineární a logitové přístupy; B) procedury založené na ortogonálních polynomech: model Martin-Bradleyův, Kronmal-Ott-Tarterův model; C) procedury založené na Matusitově vzdálenosti nominálních distribucí.

Všechny metody jsou ilustrovány příklady z aplikační praxe, jsou v některých případech i mezi sebou srovnávány co do efektivity.

Kapitola 3, *Poměry chyb a problém vychýlení*, se zabývá charakteristikami kvality rozlišení, optimalitou metody. Obsahuje jejich teoretické zavedení a metody odhadu.

Pro aplikace je velice závažná kapitola 4, *Problém výběru proměnných*, která nejprve sumarizuje přístupy lineárního klasického modelu (použití t-testu pro jednotlivé proměnné, použití standardizovaných diskriminačních koeficientů a sekvenční výběr proměnných) a ukazuje, že u *diskrétních dat mohou být hlediska na výběr proměnných zcela opačná k případu normálního rozložení*. Tyto teoretické výsledky dále narušují klid zasvěceného uživatele diskriminační analýzy, má-li k dispozici diskrétní proměnné.

Kapitola uvádí a) proceduru založenou na rozdílu diskriminačních koeficientů, b) dvě Raiffovy bayesovské metody, c) Lachinovu proceduru, d) Goldstein-Dillonovu proceduru využívající Kullbackovu divergenci. Všechny postupy jsou opět ilustrovány na reálném datovém souboru společenskovedního výzkumu (spotřeba služeb).

Kapitola 5, *Speciální témata*, je věnována důležitým oblastem, které však svým rozsahem netvoří samostatné kapitoly. Jsou to: a) případ, kdy indikátory jsou diskrétní i spojitě; b) metody pro porovnání jednotlivých technik, c) shrnutí publikovaných simulačních experimentů. Simulační studie potvrzují výhrady k lineárně klasické metodě, použité na diskrétní data: jsou-li průměry podobné, rozlišení je velmi nekvalitní, korelace proměnných podstatně snižuje přesnost klasifikace. Za-

(1) Pro srovnání uveďme, že pro k proměnných je u Fisherova modelu zapotřebí odhadnout $k(k+3)/2$ parametrů pro každou skupinu, zatímco u úplného multinomického modelu je to $m_1 \cdot m_2 \cdot m_3 \dots m_k - 1$ parametrů, kde m_k je počet hodnot k -tého znaku.