

respondentů) a nedostatky pracovních podmínek, včetně nadměrného pracovního zatížení (11,3 % dotázaných).

Z životních podmínek ve městě je nejvýše oceňována vysoká kulturní úroveň života a široké možnosti využití volného času (32,8 % dotázaných) dále úroveň občanského vybavení měst (16,1 % respondentů) a dobré zásobování (17,2 % dotázaných).

Jako nepříznivé podmínky života ve městě jsou uváděny zejména: úspěchanost a přelidněnost (36,5 %) a nepříznivé přírodní podmínky — prašnost, plynné zplodiny v ovzduší, nedostatek zeleně (24,4 % dotázaných).

Z provedeného průzkumu vyplývá také poznatek, že příznivější hodnocení podmínek života ve městě než na venkově je nepřímo úměrné stupni urbanizovanosti venkovského sídla: čím je sídlo urbanizovanější, tím menší podíl jeho obyvatel považuje životní podmínky ve městě za příznivější než na venkově.

Zajímavé je zjištění, že zájem o migraci obyvatelstva z venkova do měst během deseti let (1967—1977) poklesl téměř o polovinu. Struktura potenciálních migrantů se však v podstatě nezměnila: největší zájem o přestěhování do města mají obyvatelé neurbanizovaných sídel a mezi potenciálními migranty stále převažují mladí lidé a obyvatelé s vyšším vzděláním. K hlavním motivům potencionální migrace z venkova do měst patří i nadále především méně příznivé pracovní podmínky (včetně nadměrného pracovního zatížení a obtížnosti získat na venkově vhodné zaměstnání), nižší odměna za práci, dále celková nespokojenost s životními podmínkami, kritické hodnocení způsobu života na vesnici a rovněž nedostatečná možnost získat vyšší vzdělání a zvyšovat si kvalifikaci.

V krátkém závěru shrnuje L. V. Korelová získané poznatky a vyvozuje z nich názor na směry dalšího vývoje migrace mezi venkovem a městem v RSFSR. Příčiny snížení migrace venkovského obyvatelstva do měst stěhováním rozděluje do dvou základních skupin: do první zahrnuje příčiny, které vedly k poklesu atraktivnosti měst, a do druhé ty, které zvýšily přitažlivost venkova. V prvním případě sehrála značnou roli nasycenost měst pracovními silami a do určité míry i některé nedostatky v zásobování měst potravinami — zejména masem. Atraktivnost venkovských sídel naopak vzrostla v důsledku zlepšení jejich vybavenosti, zvýšením příjmů venkovského obyvatelstva, růstu objemu bytové výstavby, zprůměrnění práce aj.

V závěru knihy autorka prokazuje, že další vývoj stěhování obyvatelstva mezi městem a venkovem lze cílevědomě usměrňovat. Za hlavní prostředek snížení dosavadního poměrně vysokého migračního salda ve prospěch měst považuje L. V. Korelová „intenzivní urbanizaci venkovského prostoru, a to takovou, která by zachovala pozitivní stránky venkovského způsobu života, vysoce hodnocené nejen obyvateli venkova, ale též města“.

Předností knihy je účelné spojení a vyváženost teoreticko-metodologického přístupu a části analytické, která se opírá nejen o statistické údaje, ale také o výsledky vlastního výzkumu. Práce má logickou strukturu, je napsána velmi srozumitelně a přesvědčivě. Sympatické je také, že autorka dobře zná a hodnotí názory ostatních sovětských vědců a také některých odborníků z jiných zemí, zejména z USA a Polska.

Zajímavé je i obsahové uspořádání knihy. L. V. Korelová vyslovuje v teoretické části základní hypotézy o závislosti intenzity a směrů migrace mezi městem a venkovem na jednotlivých faktorech a pak tyto hypotézy na příkladu RSFSR a Západní Sibíře ověřuje.

Pro teorii marxisticko-leninské sociologie jsou přínosem zejména metodologický přístup autorky a získané poznatky o příčinách stěhování obyvatelstva z venkova do měst, a naopak. Práce je také dalším důkazem trvalé platnosti marxistického důrazu na sblížování města a venkova jako nezbytného předpokladu řešení bytové otázky ve městech a zachování relativně rovnoměrného, vyváženého osídlení celé země.

Pro sociologickou praxi bude jistě zajímavé podrobněji se seznámit s metodikou provedeného sociologického výzkumu a způsobem jeho vyhodnocení.

Souhrnně lze říci, že kniha L. V. Korelové je nesporným přínosem pro teorii i praxi marxisticko-leninské sociologie. Poučení a podněty v ní najdou kromě sociologů také demografové a pracovníci v oboru statistiky a plánování.

Alois Slepíčka

Shizuhiko Nishisato: Analysis of Categorical Data: Dual Scaling and Its Applications
(Analýza kategorizovaných dat: duální škálování a jeho aplikace)

Toronto, University of Toronto Press 1980, XIII + 276 s.

Metoda, kterou autor nazývá duálním škálováním, má asi padesátiletou historii. Její základy (teorie charakteristických čísel) však byly položeny daleko dříve Eule-

rem, Cauchym, Jacobim, Sylvestrem, Cayleyem a dalšími. První aplikace ve statistice provedl K. Pearson v r. 1901. Vlastní rozvoj statistické metodiky začíná v r. 1933, kdy H. Hotelling formuloval metodu hlavních komponent. Moderní prúdový rozvoj trvá asi patnáct let a jeho počátek se datuje publikacemi francouzského statistika J. P. Benzecriho. Metoda patří mezi explorační postupy, které se používají k odhalení vnitřní struktury matice dat nebo nějaké odvozené souhrnné tabulky za účelem hledání zdrojů vzájemných vztahů mezi řádky a sloupci tabulky.

V literatuře (především časopisecké), lze nalézt mnoho různých variant a větví metody, které vznikaly nezávisle z různých východisek, a teprve později se zjistilo, že patří k jedné velké rodině. Vyskytují se pod nejrozličnějšími názvy: metoda recipročních průměrů, aditivní skórování, vhodné skórování, kanonické skórování, Guttmanovo vážení, metoda hlavních komponent kvalitativních dat, optimální škálování, Hayashiho teorie kvantifikace, simultánní lineární regrese, faktorová analýza korespondencí, korespondenční faktorová analýza, korespondenční analýza, „biplot“, FANOVA; u nás se užívá také: metoda kanonické dekompozice, lineární dekompoziční analýza (LINDA), faktorová analýza kontingenčních (resp. profilových) tabulek, faktorová analýza reziduí.

Většina těchto názvů vznikla teoreticky při formulování různých modelů a přístupů pro řešení úloh s různými typy dat. Autor knihy používá zásadně názvu duální škálování, aby se vyhnul nedorozumění a vyjádřil jeden ze základních principů teorie. Do jaké míry se tento nový název prosadí, ukáže budoucnost. Pro českou terminologii se zdá být vhodnější termín korespondenční analýza nebo faktorová (kanonická) analýza reziduí.

V úvodu knihy autor uvádí *historii různých větví metody* a přibližuje stručně jejich východiska. Připomeňme jména některých dalších vědců, kteří se zasloužili o rozvoj metodiky: C. Eckart a G. Young (1936), P. Horst (1935), M. Richardson a G. F. Kuder (1933), H. O. Hirschfeld (1935), K. Maung (1941), L. Guttman (1941 a dále), C. I. Mosier (1946). Obzvláštní přínos znamenají publikace francouzské statistické školy koncem šedesátých a v sedmdesátých letech. Angažuje se též celá řada japonských statistiků a vytváří směr založený na pracích C. Hayashiho (od r. 1950). Sam autor knihy patří k nejznámějším specialistům v tomto oboru a vyvinul několik velmi užitečných metod.

Velmi cenným přínosem Nishisatovy knihy je to, že sjednocuje některé přístupy, ukazuje jejich vzájemnou propojenost a předkládá jednotnou teorii. Přitom se soustřeďuje na kategorizovaná data, jejich různé formy a na data pořadová a preferenční.

Principy duálního škálování (v tomto textu přijmeme terminologii autora) spolu s rozvinutím jeho matematických základů, vytvořením modelu i jeho řešením s ilustracemi na jednoduchých numerických příkladech, najdeme v *kapitolách 2 a 3*. Čtenář se tu podrobně seznámí s modelem, vycházejícím z úlohy přiřadit sloupcům kontingenční tabulky takové skóry, které by maximalizovaly koeficient determinace řádkové proměnné na takto zavedenou kardinální proměnnou i s duálními vlastnostmi opačné úlohy obdobné kvantifikace řádků. Dále jsou zde formalizovány některé další modely: model recipročních průměrů, dvourozměrné korelace, kanonických korelací a model simultánních lineárních regresí řádkové proměnné na sloupcovou a opačně. Je tu též algoritmus numerického řešení úlohy s různými doplňkovými aspekty, jako například vyloučení triviálních řešení nezávislých na datech. Důležitou částí je popis statistických testů významnosti jednotlivých kvantifikací (škál).

Ve *čtvrté kapitole Kontingenční tabulky (se dvěma vstupy) a tabulky rozdělení četností* nalezneme diskusi obou uvedených případů. Je tu ukázáno, jak souvisí duální škálování kontingenční tabulky se škálováním matice dat pro jednotlivce (tabulka vzorců odpovědí), v níž řádky odpovídají jednotlivcům a sloupce alternativám obou proměnných (řádkové a sloupcové). Postupy jsou bohatě ilustrovány (jako ostatně v celé knize) numerickými příklady.

Pátá kapitola Tabulka vzorců odpovědi: data z vícenásobného výběru (n položek); pojednává o analýze dat, která vzniknou tak, že respondent vybírá vždy jednu alternativu pro každou z n položek (kategorizovaných nominálních proměnných). Tabulka je tvořena tolika řádky, kolik je respondentů a tolika sloupce, kolik je dohromady alternativ (kategorií) u všech proměnných, jedničkou je označena volba alternativy, ostatní hodnoty v tabulce jsou nuly (součet jedniček v každém řádku je roven počtu položek). Analýza se tu komplikuje, neboť úlohu škálování lze vztáhnout k položkám i k alternativám odpovědi.

Šestá kapitola Tabulky pořadí a párových srovnání ukazuje možnosti, jak zpracovávat uvedené typy dat metodikou, která u nás není běžně známa a která je nejen

paralelou ke známým metodám (model normálního resp. logistického rozdělení), ale znamená též rozšíření na mnohorozměrnou analýzu a odhalení vícerozměrné struktury dat. Párové srovnání je řešeno metodami Guttmanovou, Nishisatovou, Slaterovou a Tucker-Carrolovou. Tyto metody jsou rozebírány a porovnány, je diskutován problém chybějících údajů a nerozlišených preferencí. Tabulka obsahující informace o pořadí je uspořádána dvěma způsoby: a) řádky = respondenti, sloupce = řazené objekty, hodnoty = centrovaná pořadí přiřazovaná respondentem, b) řádky = řazené objekty, sloupce = pořadí (1, 2, ...), hodnoty = četnosti, s nimiž se objekty umístily na pořadích.

V kapitole *sedmé Mnohorozměrné tabulky* se čtenář seznámí s problematikou více-rozměrných tabulek, které jsou prezentovány v matici, tj. jako tabulka dvourozměrná. Takováto forma tabulky však odpovídá určitému analytickému modelu a na něm pak závisí interpretace. Kapitola neobsahuje obecné řešení úlohy, ale poskytuje „pracovní znalost“ (jak říká autor) odvozenou z různých případů. Optimální škálování se zde může týkat různých hledisek, a proto se model formuluje pro každý případ zvlášť.

V *osmé kapitole Různá témata* je nejdůležitější problém škálování, při němž se požaduje zachování konzistence skóre s předem daným pořadím kategorií. Dále je tu rozebrána aplikace duálního škálování na řešení úlohy analýzy rozptylu pro kategorizovaná data a problém chybějících údajů.

V *příloze A* jsou shrnuty potřebné znalosti z teorie matic. Pro praxi teorie duálního škálování je významný *dodatek B*, v němž je výpis fortranského programu založeného na metodách knihy. Program je doplněn testovacími příklady, které prověřují různé větve výpočtů a které podstatně ulehčí implementaci programu pro počítač. Případné odchylky od použité varianty jazyka FORTRAN opraví každý kvalifikovaný programátor. Obdobně není obtížné upravit dimenze polí, pokud možné rozměry tabulek určené v programu uživateli nevyhovují. Metody knihy, které nelze řešit bez pomoci počítače, jsou tak k dispozici každému zájemci. Seznam odkazů je velmi široký a slouží jako vodítko k dalšímu podrobnému studiu.

S. Nishisato klade důraz na optimální kvantifikaci kategorií, což zdůrazňuje i názvem „duální škálování“. Pro využití v sociologickém výzkumu se zdá být vhodnější faktorový model, který se objevil v publikacích v Sociologickém časopise v poslední době. Jsou-li jednotlivé stupnice chápány jako projevy latentních faktorů, mohou interpretace vést k hlubšímu pochopení vzniku variability dat a mohou sloužit k hledání vnitřních příčin, které ovlivňují vznik a seskupení dat v dané tabulce. Druhé, u nás též doporučené a v praxi značně populární využití je grafické zobrazení vztahů, které se v tabulce projevují. Geometrická reprezentace řádků a sloupců tabulky vychází z přístupu mnohorozměrného škálování korespondencí. Faktorová interpretace modelu odpovídá snaze nalézt podstatné vlivy, příčiny, společné zdroje rozdílnosti a podobnosti. Smysl geometrického zobrazení je v převodu složitých vztahů v datech do přehledné a názorné formy. Oba způsoby mají silný vliv na sociologickou imaginaci, na vytváření zcela nových hypotéz, na rychlé odhalení netypických jevů, empirických pravidelností a i jen náznakových tendencí a vztahů. Mohou tak zachytit jevy v jejich zárodečných stádiích. Tyto dvě možnosti autor ve svém výkladu pomíjí, to však neznamená, že výsledky teorie nelze uvedeným způsobem interpretovat. Jak bylo řečeno výše, stejná teorie vychází z celé řady výchozích modelů, takže číselné výsledky programu DUAL 1 z dodatku lze chápat jako faktorové koeficienty nebo jako souřadnice bodů v eukleidovském prostoru.

Nishisatova kniha znamená značný přínos pro praxi analýzy dat i pro popularizaci a rozšíření metody. Text je vysoce profesionální, dobře strukturovaný podle stupně obecnosti a podle typů dat, výběr postupů je konzistentní, komentáře jsou jasné, promyšlené a velmi instruktivní, příklady jsou vybrány tak, aby ilustrovaly širokou variabilitu situací a možností. Je to monografie, která sice může působit dojmem jisté jednostrannosti (překonatelné, jak bylo naznačeno výše), která však téma pojednává do hloubky a vhodně pro praktické použití. Měl by se s ní seznámit každý, kdo má metodologické ambice. Metoda by měla být též přednášena studentům a aspirantům, neboť trendy v analýze kategorizovaných dat ukazují, že je to metoda budoucnosti, bez níž se nová generace výzkumníků ve společenských vědách zřejmě neobejde.

Blanka Řeháková