

8. Uskutečnit pracovní jednání skupiny expertů, které by zvažilo další možnosti spolupráce v oblasti sociologického výzkumu ateismu a náboženství.

Ivan Tomek, Ján Mišovič

Mezinárodní setkání odborníků na analýzu společenskovědních dat

Počátkem října 1986 se na ostrově Hiddensee (NDR) uskutečnil již čtvrtý ročník mezinárodního semináře *Strategie statistické analýzy dat*, který pořádalo Centrum pro společenskovědní informace Akademie věd NDR spolu s Univerzitou Ernsta Moritze Arndta v Greifswaldu.

Základním cílem semináře bylo, stejně jako v minulých letech, dát prostor pro výměnu zkušeností z oblasti statistické analýzy dat ve společenských vědách a seznámit účastníky s novinkami v oblasti programového vybavení pro potřeby společenských věd. Značná pozornost byla věnována také využívání výpočetní techniky a praktickým zkušenostem s aplikací této techniky v empirických výzkumech.

Důležité a velmi podnětné je, že na semináři se setkávají tvůrci programového vybavení — matematikové, statistikové a programátoři — s jeho uživateli. Je ovšem příznačné, že obě skupiny odborníků si velmi dobře rozumějí, neboť k účasti jsou zváni pracovníci s dlouholetou praxí v aplikaci výpočetní techniky a statistických metod. Interdisciplinarity semináře je navíc zdůrazněna tím, že se na něm již tradičně u jednoho stolu setkávají specialisté působící v různých společenských, ale i přírodních vědách — sociologii, psychologii, ekonomii, historii, geografii, geologii apod. Poznatky z využití různých postupů analýzy dat v těchto na první pohled velmi odlehklých oborech jsou velmi inspirativní.

Zatímco v letech 1983 a 84 byla hlavní pozornost zaměřena na „zmapování“ existujícího programového vybavení vhodného pro využití ve společenských vědách a na výměnu zkušeností s jeho aplikací, v loňském i letošním roce byl důraz položen do některých speciálních oblastí zpracování dat a do teoretické sféry. Těžištěm se v obou posledních ročnících stala problematika strukturních vztahů a modelů lineárních strukturních rovnic. Tato metoda práce s daty, která má hluboké metodologické důsledky, se v současné době dosti dynamicky rozvíjí a nachází stále širší uplatnění v řadě společenských věd, zejména v ekonometrii, sociologii a psychologii. Zjednodušeně řečeno, jde o soubor modelů analýzy kovariačních struktur, s jejichž pomocí je možné do určité míry odhalit vztahy ve složitých systémech, analyzovat směr a sílu působení závislostí jak mezi pozorovanými indikátory, tak mezi hypotetickými latentními (nepozorovatelnými) proměnnými. Zmíněné metody jsou spojeny zejména se jmény K. Jöreskoga a R. P. McDonalda, kteří počátkem sedmdesátých let poprvé zformulovali základy modelů strukturních rovnic. V současné době existuje kromě nejznámějšího Jöreskogova modelu řada dalších, které jsou v některých směrech i obecnější.

Různé modely analýzy kovariačních struktur byly základem pro vytvoření několika programových systémů, z nichž nejstarší, nejrozšířenější a nejznámější je systém LISREL (Linear Structural RELations), který je v současné době k dispozici již v šesté doplněné verzi a o němž se lze dočíst v uživatelské příručce Jöreskog, K. G. — Sörbom, D.: LISREL VI. — Analysis of Linear Structural Relationships by the Method of Maximum Likelihood (Scientific Software Inc., Mooresville 1985). O teoretických základech, na nichž je tento programový systém vybudován, o možných variantách aplikací na různé datové struktury a zkušenosti s jeho používáním referoval prof. Schmidt (Univerzita Giessen). O zkušenostech s využíváním téhož programového systému v oblasti sociálních výzkumů hovořil F. Faulbaum (Zentrum für Umfragen, Methoden und Analysen, Mannheim), který se zaměřil na použití uvedených metod pro analýzu vlivu panelového efektu v longitudinálních výzkumech mínění při sledování názorů jednotlivých populačních kohort v rámci určitého časového intervalu. Použití bylo demonstrováno na příkladech výzkumu vztahu ke „gastarbeitern“ v NSR.

Model vycházející z obdobných metodologických přístupů k analýze dat představil na semináři prof. McArdle (University of Virginia), který zpracoval také odpovídající programový produkt RAM. Podrobnosti o tomto modelu obsahuje článek McArdle, J.: *Causal modelling applied to psychonomic systems simulation*. Behav. Res. Meth. Instrument., 12, 1980, s. 193 — 209. a za pozornost jistě stojí, že program RAM má být v dohledné době zařazen do jednoho z nejrozvinutějších komerčních systémů pro statistickou analýzu dat — SAS.

Z poněkud jiného modelu analýzy kovariačních struktur vychází programový systém PLS (Latent Variables Path Analysis with Partial Least Squares), který byl navržen a do programové podoby realizován E. Löhmoellerem (Freie Universität,

Berlín). Ten nejprve seznámil účastníky semináře s teoretickými základy svého modelu a poté na řadě příkladů z výzkumné praxe demonstroval různé možnosti využívání jeho jednotlivých komponent při řešení různých výzkumných úloh. Zajímavá byla i druhá část vystoupení, kde byly vzájemně porovnány možnosti, výhody, limity, ale i výpočetní rychlost a schopnosti různých programových systémů, realizujících jednotlivé modely analýzy kovariačních struktur. Kromě již zmíněných systémů LISREL, RAM a PLS se jedná o program EQS, pracující v rámci systému BMDP (bližší o něm viz Bentler P. M.: *Theory and implementation of EQS, a structural equation program*. BMDP statistical software, Los Angeles 1983) a systém COSAN (ten je popsán v článku McDonald R. P.: *A simple comprehensive model for the analysis of covariance structures: Some remarks on applications*, Brit. math. statist. psychol., 33, 1980, s. 59–72). Srovnání ukázalo, že v současné době musí uživatel, který uvažuje o nákupu některého z těchto softwareových produktů, zvážit, zda dává přednost systémům bohatším, ale též dražším, nebo jednodušším a levnějším, které jsou také při provozu méně náročné na strojový čas a paměť počítače. Všechny představené programy jsou, nebo v nejbližší době budou, k dispozici jak ve verzi určené pro střední a velké počítače, tak v podobě vyhovující provozu na osobních počítačích.

Druhým okruhem otázek řešených na semináři byly problémy vytváření informačních systémů v oblasti společenských věd a možnosti integrace těchto systémů se statistickou analýzou dat. V této části byly předneseny dva základní příspěvky.

T. Cierzynski z Matematického ústavu AVNDR v Berlíně informoval o nově vytvořeném informačním systému, který je součástí budované banky programů. Cílem snažení ústavu je získat tímto způsobem přehled o existujícím software, integrovat dokumentaci o programech a konečně ovlivňovat další práce na vytváření nových programů. Informační systém umožňuje prostřednictvím databázové technologie získávat různé informace jak pro odborníky v jednotlivých vědních oborech o programových možnostech řešení jejich problémů, tak pro pracovníky výpočetních středisek o dislokaci požadovaných programových produktů a konečně pro matematiky též informace o realizovaných metodách nebo rešerše literatury týkající se dané problematiky.

Druhý příspěvek přednesl J. Řehák (Geografický ústav ČSAV, Praha), který podobně informoval o koncepci zajímavého informačního komplexu pro uchovávání geografických dat, charakterizujících různé stránky jednotlivých regionů, pro jejich neustálou aktualizaci s možností statistického zpracování prostřednictvím odpovídajícího programového vybavení. S realizací tohoto komplexu se právě začíná a jeho výstavba je rozvržena do několika etap, v nichž budou postupně vytvořeny jednotlivé komponenty informačního systému — banka dat, banka programů a metodologická banka.

Zajímavá je idea uchovávat a zpracovávat nejen „klasická“ číselná data, ale i grafickou geografickou informaci. Na koncepci je pozoruhodná zejména komplexnost v přístupu k problému. Nezůstává se pouze u uchovávání a zpracování dat, ale pozornost je věnována také metodologickým problémům, bude vytvořen dokonce expertní systém, sloužící pro volbu optimální koncepce zpracování datových souborů.

Mimo hlavní dva problémové okruhy odeznělo na semináři několik dalších zajímavých vystoupení. M. Čákr (Výzkumný ústav strojírenské technologie a ekonomiky, Praha) referoval o problematice práce s chybějícími daty v sociologickém výzkumu, J. Herzmann (ÚVVM, Praha) o zkušenostech s využíváním načítaných skóru při analýze výzkumů veřejného mínění, D. Bellachová (Humboldtova Univerzita, Berlín) o převedení problému volby koeficientu souvislosti proměnných do jazyka teorie rozhodovacích funkcí, K. Formella (Univerzita Greifswald) o programovém zabezpečení zpracování dat v univerzitním výpočetním středisku, D. Pokorný (Psychiatrický ústav ČSAV) o některých psychologických aspektech používání faktorové analýzy.

Kladem celého semináře bylo, že v místnostech, v nichž jednání probíhala, byly instalovány dva osobní počítače, na nichž bylo přímo demonstrováno zpracování dat vztahujících se k příkladům uváděným v diskusních příspěvcích.

Mezinárodní seminář přinesl všem účastníkům řadu konkrétních podnětů pro práci s daty a ukázal další možnosti využívání výpočetní techniky. Jednoznačně se projevil trend přechodu ke komplexním metodám analýzy dat, z nichž právě lineární strukturní modely zaznamenávají v současné době nejprudší rozvoj. Současně byla patrná stále větší orientace na využívání osobních počítačů, které nyní umožňují použití všech běžně používaných statistických metod i práci s rozsáhlými datovými soubory. Účastníci se shodli na tom, že slibně se rozvíjející dialog by měl v budoucnosti pokračovat za účasti pokud možno ještě širšího okruhu předních specialistů, kteří se strategickými otázkami analýzy společenskovědních dat zabývají.

Václav Forst, Jan Herzmann