

Foret, Miroslav; Librová, Hana

Program SOC, jeho možnosti, meze a dosavadní zkušenosti

Sborník prací Filozofické fakulty brněnské univerzity. G, Řada sociálněvědná.
1973, vol. 22, iss. G17, pp. 129-133

Stable URL (handle): <https://hdl.handle.net/11222.digilib/111387>

Access Date: 02. 12. 2024

Version: 20220831

Terms of use: Digital Library of the Faculty of Arts, Masaryk University provides access to digitized documents strictly for personal use, unless otherwise specified.

Samočinné počítače používáme při zpracování dat proto, aby nám ušetřily práci, čas a náklady. Příprava programů pro analýzy výsledků sociologických výzkumů a sociálních ukazatelů a dat je záležitost pracná, nákladná a náročná na čas. Tak například na systému programů SPSS, který je užíván v řadě států a nedávno byl zakoupen i pro některé počítače v Československu, pracovalo deset odborníků a jeho dovedení do konečné podoby trvalo téměř deset let. Výsledkem je velmi variabilní sada programů s širokým spektrem možností, pracující opravdu efektivně. Nevyplatí se nám ale vždycky zakoupit tak rozsáhlý programový systém, je velice drahý. I při tvorbě specializovaných systémů s užším rejstříkem možností však musíme bedlivě sledovat otázku jejich ekonomiky, a to nejenom ekonomiky tvorby programů, ale i jejich výsledné parametry z hlediska efektivnosti při zadávání, zpracování a vyhodnocení dat. Všechny tyto aspekty jsou u každého, i relativně jednoduchého systému ve vzájemné závislosti a složitě se ovlivňují. Konečné cílevědomé vyvážení má pak rozhodující vliv na produktivitu systému a mnohonásobně se vyplatí. Bylo by chybou tuto skutečnost přehlížet. Trvalý zřetel na ekonomiku není nedůstojný vědy. Volba přiměřené ekonomické strategie se stává neodlučnou součástí každého výzkumného projektu, programování pro samočinné počítače nevyjímaje, pochopitelně.

Ivo Možný

PROGRAM SOC, JEHO MOŽNOSTI, MEZE A DOSAVADNÍ ZKUŠENOSTI

Již více než čtyři roky mají zaměstnanci a studenti oddělení sociologie filozofické fakulty UJEP možnost využívat vlastní program na statistické zpracování empirických údajů — program SOC. Celý program vznikl ve Vědeckometodickém středisku pro výpočetní techniku při katedře aplikované matematiky na přírodovědecké fakultě UJEP na tamním samočinném počítači MSP-2A.¹

V době, kdy jsme vypracovali program SOC nebylo technické vybavení počítače MSP-2A příliš vhodné pro užití při zpracování materiálů z empirických sociologických výzkumů, které klade vysoké požadavky zejména na vstupní a výstupní zařízení samočinného počítače a na kapacitu jeho paměti. A právě tyto prvky patří na počítači MSP-2A mezi slabé. Proto jsme na návrh M. Gregora omezili formu vstupních dat na *data výhradně předem vytříděná*.

Druhou odlišností programu SOC oproti ostatním programům používaným ke zpracování dat ze sociologických výzkumů na samočinných počítačích je jeho obsah. Program SOC se zaměřuje na stále ještě nejčastější případ znaků, které se v současných sociologických výzkumech vyskytují — tj. na *znaky nominální a ordinální*.

Konečně třetí koncepční zvláštností programu SOC měla být *jednoduchost vstupních instrukcí a přehlednost i srozumitelnost výsledků*. Program SOC by měl umožnit i sociologům neškoleným v otázkách využívání samočinných počítačů počítat některé statistické charakteristiky.

Program SOC byl vyhotoven v programovacím jazyku AUTOKÓD AU—MSP-2A, přesněji v jeho variantě AU 10.

Základní charakteristiky programu SOC:

I. Forma vstupních dat

Program SOC není, jak jsme již uvedli, v žádném případě schopen provádět třídění dat. Pro jeho vstup je třeba mít data již předem vytříděná — u malého počtu zkoumaných případů ručně; u větších na mechanických třídíčkách, tabelátorech apod., či používat již vytříděných tabulek na jiných samočinných počítačích atd.

¹ Počítač MSP-2A je československým počítačem druhé generace a koncepčně zapadá do stavu našich projektů samočinných počítačů v letech 1963—1965. Dosahuje rychlosti okolo 7000 operací za vteřinu. V době vzniku programu SOC v letech 1969—1970 byla jeho jedinou pamětí ferritová paměť s kapacitou 10 000 slov. Nejrychlejším vstupem je snímač děrné pásky typu FS 1500 s rych-

Dvourozměrná tabulka absolutních četností vytváří tzv. pásku dat. Její základní forma vypadá následovně

| | | | | | | | |
|----------|----------|---|---|---|---|---|----------|
| n | m | | | | | | |
| x | y | | | | | | |
| a_{11} | a_{12} | . | . | . | . | . | a_{1m} |
| a_{21} | | | | | | | . |
| . | | | | | | | . |
| . | | | | | | | . |
| . | | | | | | | . |
| a_{m1} | a_{m2} | . | . | . | . | . | a_{mn} |
| 9999 | | | | | | | |

kde n je počet sloupců počítané tabulky, m je počet řádků, x je číselné označení položky v sloupcích, y je číselné označení položky v řádcích a a_{11} , a_{12} , a_{13} . . . , a_{mn} jsou číselné hodnoty absolutních četností na jednotlivých políčkách načítané tabulky.

Každá vstupní tabulka musí nutně zahrnovat všechny tyto údaje. Nevíme-li například, jaké číselné hodnoty náleží x a y , musíme si je zvolit alespoň ad hoc. Na konci každé tabulky musí být nezbytně uveden kontrolní znak 9999.

Mezi číselné hodnoty na prvním a druhém řádku předpisované tabulky (tj. mezi řádek končící m a řádek začínající x) lze vsunout řádek s libovolně dlouhým textem, který se vytiskne v záhlaví výsledné tabulky a jehož vzorový předpis vypadá takto: „ = LIBOVOLNE DLOUHY TEXT blanky a řádek“.

Pro výpočet průměru však musí páska dat ještě obsahovat další nezbytné hodnoty – „váhy“. Je-li za zakončenou tabulkou (tj. za 9999) číslo 1313 (resp. 1414), generuje program tolik čísel ze vzestupné posloupnosti přirozených čísel 1, 2, 3, . . . , kolik má počítaná tabulka řádků (resp. sloupců). Jsou-li však za tabulkou uvedeny „váhy“, načte se jich vždy jen tolik, po kolik řádků nebo slouců (resp. obojí) se budou průměry počítat. Potom by mělo následovat načítání další tabulky absolutních četností nebo celkový konec výpočtu.

Program SOC je schopen zpracovat tabulku absolutních četností o maximální velikosti $m = 30$ a $n = 20$. Četnost tabulky v celkovém součtu nesmí přesáhnout 99 999. Počet počítaných tabulek není nijak omezen. Při větším souboru stejně zpracovávaných tabulek je výhodné za poslední tabulkou vyděrovat navíc 1010 a provádět celý výpočet pomocí tzv. „nekonečné řídicí pásky“.

II. Obsah programu SOC

Formálně lze program SOC považovat za systémovou stavebnici sestávající z těchto devatenácti částí:

1. Načítání vstupních dat, jejich kontrola (je-li cekový součet absolutních četností v tabulce roven nule, vypíše se na dálnopise NESMYSLNA DATA a výpočet se zastopuje, není-li na $(m \times n) + 1$ —ém místě po zahájení načítání absolutních četností kontrolní znak 9999, vypíše se na dálnopise CHYBA V PASCE DAT a výpočet se opět zastopuje atd.) a počítání marginálních četností tabulky. Zároveň se tiskne hlavička výsledné výstupní tabulky.
2. Tiskne se tabulka *absolutních četností* s okrajovými součty a součtem celkovým. Jednotlivé hodnoty se tisknou na maximálně pět platných míst, tj. nejvyšší celkový součet v tabulce může být 99 999.
3. Počítají se a tisknou *relativní četnosti řádků* a sloupců. Tisk všech relativních četností je na jedno desetinné místo.

lostí 1500 znaků/sec. Nejvýhodnějším výstupem v té době byl děrovač děrné pásky (typ PE Facit); proto bylo tehdy třeba výsledné děrné pásky nechat vypisovat na elektrickém psacím stroji Consul, doplněním snímačem děrné pásky. Snad jedinou výhodou výstupu na děrovači děrné pásky je možnost jednoduchého použití nepoškozené děrné pásky s výsledky k mnohonásobnému vytištění výsledných tabulek či jako vhodného operativního způsobu uložení výsledků.

4. Vypočítávají se a tisknou **relativní četnosti sloupců** a procenta marginálních četností v řádcích. Tisk je opět na jedno desetinné místo.
5. Vypočítává a tiskne ve tvaru tři místa před jedno za desetinnou tečkou očekávané četnosti. Na okrajích tabulky se tisknou řádkové a sloupcové chí-kvadráty s kritickými tabulkovými hodnotami na hladinách významnosti 10 %, 5 %, 1 % a 0,1 %.
6. Vypočítají se a na dvě desetinná místa vytisknou korelace alternativ.
7. Počítá a v příslušné formě vytiskne testovací kritérium znaménkový test.
8. Počítá a vytiskne chí-kvadrát na pět platných míst před a na čtyři za desetinnou tečkou. Tiskne se i příslušná kritická hodnota a její hladina významnosti.
9. Vypočítá a na čtyři desetinná místa vytiskne Čuprovův koeficient kontingence.
10. Počítá a tiskne Pearsonův koeficient kontingence a normovaný Pearsonův koeficient kontingence.
11. Počítá a vytiskne na čtyři desetinná místa Spearmanův koeficient pořadové korelace. Tisknou se i kritické hodnoty, získané z tabulek pro koeficient korelace.
12. Vypočítává se a na čtyři desetinná místa se tiskne normovaný Pearsonův koeficient kontingence podle Hofstättera.
13. Počítá relativní četnosti a průměry řádků, které stejně jako u následujícího výpočtu se tisknou ve tvaru tři číslice před a jedna za desetinnou tečkou. Ve stejném tvaru se na místo označení sloupců tisknou „váhy“.
14. Vypočítává relativní četnosti a průměry sloupců a tiskne je.
15. Provádí výpočet a vytisknutí celkových relativních četností.
16. Střídavé určování závisle a nezávisle proměnné v tabulce vede ke dvěma výsledkům Goodman-Kruskalovy lambdy, které se tisknou na čtyři desetinná místa.
17. Počítá a vytiskne na čtyři desetinná místa Goodman-Kruskalovu gamu.
18. Vypočítá a vytiskne charakteristiky a koeficienty vycházející z teorie informace — entropií a množství informace s výběrovými chybami a koeficienty souvislosti. Testovacím kritériem pro množství informace jsou příslušné hodnoty jednostranného normovaného normálního rozložení.
19. Proveďte automatické rozhodnutí, zda se na podkladě splnění příslušných podmínek bude dále počítat a tisknout jedna ze tří uvedených větví:
 - a) Fisherův test,
 - b) oba Yuleovy koeficienty a koeficient závislosti pro čtyřpolní tabulky s výběrovými chybami,
 - c) Cramerův koeficient a Scottova míra shody.

III. Forma příkazových (řídících) instrukcí

Uvedené členění obsahu programu SOC umožňuje skutečné řízení výpočtu na samostatném počítači. Přidá-li se k číselnému označení částí programu 1, 2, 3, 19 aktivní znak — trigger, vytvoří se tzv. řídicí páska, která předává uživateli požadavky programu a která musí vždy začínat načtením vstupních dat tabulky — tj. 1. Další požadované početní postupy je však možno zadávat v libovolném pořadí, bez jakýchkoli návazností. Nejsou-li jiné požadavky, je z ekonomického hlediska výpočtu výhodnější zadávat řídicí pásku jako vzestupnou řadu. Je například hospodárnější si zadat výpočty statistických charakteristik v pořadí 1 5 8 9, než v pořadí 1 9 8 5. V druhém případě by se totiž počítaly očekávané četnosti celkem třikrát a chí-kvadrát dvakrát, zatímco v prvním případě se tyto výpočty provedou jen jednou a zůstávají uchovány v paměti počítače i pro další potřeby.

Končí-li páska dat (kterou chceme celou zpracovávat stejným způsobem pro všechny tabulky) znakem 1010, je možno vytvořit z jednoho (nebo několika málo) zadání a slepením obou konců dérné pásky tzv. „nekonečnou“ řídicí pásku. Jestliže však není páska dat zakončena 1010 (nebo nejsou-li u všech tabulek stejné požadavky na jejich zpracování), je nutné děrovat řídicí pásku pro každou zpracovávanou tabulku zvlášť a nakonec vyděrovat — aktivní znak pro zastavení výpočtu.

IV. Forma výstupních dat (výsledků)

Přes všechny technické potíže s výstupem programu SOC na samočinném počítači MSP-2A byla snaha opatřit v konečných výsledcích každou vypočtenou statistickou charakteristiku natolik jasným komentářem, aby již sociolog k jejich interpretaci nepotřeboval žádné objasňující vysvětlení, klíč apod.

Změny počítače MSP-2A a jejich vliv na program SOC

Zaměstnanci Vědeckometodického střediska pro výpočetní techniku při katedře aplikované matematiky UJEP samočinný počítač MSP-2A neustále systematicky zlepšují, a to nejen po stránce vybavení programy, operačním systémem apod. — tedy po stránce softwaru, ale i po stránce technického vybavení počítače — tedy po stránce hardwaru. Z technického hlediska byl samočinný počítač MSP-2A ve středisku od roku 1970 doplněn vnější magnetickou pamětí a širokořádkovou tiskárnou (která tiskne 140 znaků na řádek rychlostí 600 řádků za minutu).

V letech 1971–1973 došlo ve Vědeckotechnickém středisku také k přestavbě operačního systému samočinného počítače. Byl vytvořen nový operační systém, což konkrétně pro náš program SOC znamenalo, že byl v důsledku svých vysokých nároků na kapacitu vnitřní ferritové paměti provozu neschopný. Proto bylo nutné program SOC „renovovat“ — z devatenácté části byly vypuštěny Fisherův test, Cramerův koeficient a Scottova míra. Dnes tedy používaná a jediná provozuschopná verze programu SOC již tyto statistické charakteristiky nemá.

V současné době je program SOC navíc schopen využívat širokořádkovou tiskárnu, a to dvojnásobem: 1. výsledky nejdříve vycházejí jako obvykle z děrovače děrné pásky a potom se výpis děrné pásky místo na elektrickém psacím stroji provede přes samočinný počítač na širokořádkové tiskárně,

2. výstupní kanál se hardwareově přepojí z děrovače děrné pásky přímo na širokořádkovou tiskárnu.

Bohužel obojí využití širokořádkové tiskárny není nejekonomičtější, neboť aranžmá výstupních výsledků bylo vypracováno na odlišný formát papíru a na jiné posuvné mechanismy elektrického psacího stroje.

Jak tedy vidíme, vzhledem k tomu, že práce na programu SOC byly zastaveny v roce 1970, není v současné době bohužel možné například dokonale využít všech výhodných nových úprav samočinného počítače MSP-2A ve Vědeckotechnickém středisku pro výpočetní techniku UJEP.

Dosavadní zkušenosti s programem SOC

Použití programu SOC v sociologické práci je v současné době umožněno tím, že Vědeckotechnické středisko pro výpočetní techniku při Katedře aplikované matematiky uzavřelo s oddělením sociologie smlouvu o bezplatném použití počítače MSP pro pedagogické účely a pro odbornou práci oddělení sociologie. Taková smlouva a její realizace svědčí mimo jiné o tom, že spolupráce mezi fakultami University — v tomto případě mezi fakultou přírodovědeckou a filosofickou — není pouhým přáním, jemuž se formálně vyhovuje, ale že se uskutečňuje zcela konkrétně v denní praxi.

Původní záměr tvůrce programu (podotýkám zde jenom mimochodem, že program je vedlejším výsledkem diplomové práce posluchače oboru sociologie) bylo, jak je již výše zmíněno, doplňovat tabulky zpracované na jiných počítačích některými dalšími statistickými charakteristikami. Tento záměr ustupuje v současné době do pozadí a čím dál tím více se ukazuje, že program SOC je dobrým pomocníkem i pro účely pedagogické. Máme zde na mysli především práci s programem v praktických cvičeních z metod a technik sociologického výzkumu a dále využití programu posluchači při zpracování empirických částí diplomních úkolů. Program posluchačům umožňuje hlouběji pochopit a prakticky si ověřit techniku korelační analýzy sledovaných znaků.

Studenti ve cvičeních z kvantitativních metod sociologického výzkumu provádějí většinou sekundární analýzu dat z již realizovaných výzkumů. Data mají přitom k dispozici ve formě kódů na děrných štítcích. Zvolená třídění provedou posluchači mechanicky na třídičce děrných štítků, kterou disponuje oddělení sociologie. Na počítač MSP dodají potom zadání výpočtů charakteristik těsnosti a významnosti vztahů mezi vybranými znaky ve formě dvourozměrných tabulek absolutních četností.

Zde je na místě zmínit se o tom, že neschopnost programu SOC třídit data je nedostatkem spíše z hlediska technického. Pro účely pedagogické je tato „slabá stránka“ programu spíše přínosem. Posluchači, kteří s programem pracují, jsou přinuceni vytřídit data ručně nebo mechanicky. Po zkušenostech můžeme říci, že přípravné práce jsou přínosné z pedagogického či didaktického hlediska, protože distribuce děrných štítků v jednotlivých odděleních mechanické třídičky ilustruje a doslova hmatatelně demonstrovuje existenci či absenci vztahů mezi sledovanými znaky, což dále napomáhá zafixování představ o souvislosti mezi některými základními sociálními jevy a tím k celkové zběhlosti a orientaci posluchače ve studovaném oboru.

Dále: poměrná pracnost mechanického třídění nutí studenta – a to je pro jeho odbornou výchovu důležité – aby třídil a tedy i zadával na počítač jen výpočty takových vztahů, které mohou mít skutečně smysl a meritorní opodstatnění. Jinými slovy, fakt, že program SOC nutí posluchače zadávat již vytřídná data, může napomoci tomu, aby se student v pozdější praxi nedopouštěl běžné chyby, k níž jinak použití samočinných počítačů svádí. Myslíme zde právě zadávání korelací všech možných proměnných bez výběru, které potom vede k zahlcení statistickým materiálem v tabulkách.

Jak jsme se zmínili, strojového času počítače MSP pro práci s programem SOC využívají i studenti vyšších ročníků, kteří pracují na diplomních úkolech empirického charakteru. Podle zadání diplomové práce pracují i zde někteří posluchači již se sebraným statistickým materiálem, tj. provádějí pomocí programu SOC tzv. sekundární analýzu informací, které dosud nebyly vyčerpávajícím způsobem zpracovány a jsou k dispozici na oddělení. Vedle toho používají tohoto programu diplomanti i ke statistickému vyjádření souvislostí mezi jevy, které sledovali vlastním výzkumem, vlastním sběrem dat. Program SOC tak umožňuje posluchači, aby ještě před absolutoriem provedl samostatně výzkum od přípravné až po závěrečnou fázi. Vzhledem k tomu, že na velkých počítačích je strojový čas v současné době velmi vzácný, těžko dostupný a konečně drahý, je to jediná možnost, jak můžeme studentům takovou možnost běžně poskytnout.

Zmírně se ještě na závěr o tom, že program SOC slouží i pracovníkům sociologie při práci malého rozsahu, a to zejména při doplňujících analýzách již realizovaných a zpracovaných výzkumů. Program může právě z tohoto důvodu přispět k zefektivnění práce výzkumných pracovníků, neboť má schopnost pomoci tam, kde je slabé místo většiny sociologických výzkumů, jakým je sběr velkého množství informací, které bývají po příchodu z terénu nedokonale využity.

Miroslav Foret
Hana Librová

SOCIOLOGICKÁ PRÁCE V MUZEÍCH

V uplynulém desetiletí byla v řadě zemí založena muzeologická střediska (katedry na vysokých školách nebo samostatné instituty), jejichž posláním je zajistit doplňkové muzeologické vzdělání jak studentů některých oborů, tak především muzejních pracovníků. Celosvětový rozvoj muzejnictví, stále stoupající zájem veřejnosti o muzejní expozice a výstavy si vyžádaly, aby také této významné kulturní oblasti byla věnována odpovídající odborná pozornost, specializovaný teoretický výzkum. Muzejní praxe pak potvrdila, že odborníci zaměstnaní v muzeích musí ovládat nejen některou z vědních disciplín, jež jsou tak či onak dokumentovány v těchto institucích, ale že se neobejdou ani bez znalosti muzeologické teorie. Tato teorie (jak to dokládá i publikační aktivita zmíněných výukových středisek) získává stále více charakter samostatného vědního oboru – muzeologie. S upřesňováním jeho předmětu se současně zjišťují