

only what he (she) is, but what he (she) is becoming, what he (she) may become. The questionnaire must also perform the function of a "textbook", it is a means of motivating practical commitment; it not only investigates attitudes but also evokes them. Besides being a means of ascertaining the condition of the view of the world, it is a form of verifying a certain categorial system, a certain theoretical system. Closed answers and their quantifications impair and level out the originality and individuality of the examined subjects. The questionnaire, which is nothing more than a collection of closed questions, is not suitable for investigating the problems of the view of the world. The aspects of form and content of the questionnaire must respect the personality of the subject as a unique individual who cannot be reduced to an "equalized number one". It is absolutely impossible to quantify and

mathematically manipulate the subject as a unique quality.

It is difficult to interpret the data obtained through a research carried out by means of questionnaires. Every synthesis means a reduction. A more adequate interpretation presupposes repeated researches. We do not know the level of the subjects' knowledge and choice. The subjects make statements only about what they are conscious of, not about everything that affects their views. All we know or should know about the young people's view of the world cannot be communicated by the young. In investigating the political attitudes, it is not possible to abstract from the individual's institutional context. The sociological definition of the young man's (woman's) view of the world is privileging certain aspects, it is partial, it lacks the ethical, the historical, the philosophico-anthropological, the biopsychical and the physiological dimensions.

Použití informačních testů homogenity při výzkumu vysokoškolských studentů

V tomto příspěvku chceme upozornit na použití jedné metody matematické statistiky při zkoumání empirických dat sociologických výzkumů. Jde o možnost využití statistik konstruovaných na základě teorie informace pro testy homogenity statistických rozložení. Této metody bylo použito ve výzkumu vysokoškolských studentů, realizovaného v roce 1966 Ústavem marxismu-leninismu pro vysoké školy v Praze.

Základní soubor a výběrový soubor

Výběrový soubor, jehož elementy byly objektem výše uvedeného sociologického výzkumu, byl tvořen 2726 posluchači 1. ročníků devatenácti záměrně vybraných fakult vysokých škol. (Základní soubor = 23 539.) Analýza předcházející volbě charakteru výběrového souboru i charakteru výběrového řízení vyústila v závěr (detailně popsán v projektu výzkumu), že není vhodné ani účelné uvažovat o realizaci výzkumu na celém základním souboru, ani na jeho části, vybrané např. prostým náhodným výběrem, že není účelné brát v úvahu všechny drobné skupiny podle lokalit i studijních specializací. To vedlo k záměrnému výběru podle

dvou hlavních diferenčních kritérií — typ fakulty, lokalita. Výběr se opíral o dvě základní hypotézy:

a) V určitých elementech existují významné rozdíly mezi studenty fakult rozdílného odborného zaměření.

b) V některých elementech existují rozdíly mezi studenty studujícími na vysokých školách, vybudovaných po roce 1945 a umístěných v malých, z daného aspektu marginalních lokalitách, a mezi studenty studujícími na vysokých školách s dlouhodobou tradicí, umístěných v centrálních městech s rušným akademickým i vědeckým životem, v lokalitách, jež jsou centrem politického, hospodářského i kulturního života.

Tyto výchozí hypotézy byly determinantou voleb jednotlivých typů fakult a lokalit, a to tak, aby bylo možno provést srovnání v obojím směru.

Tak bylo záměrně vybráno devatenáct fakult pěti typů odborného zaměření (z celkového počtu 89 fakult 27 vysokých škol v ČSSR). Ve vybraných fakultách (viz v přehledu v tabulce 1) bylo náhodně vybráno dohromady 2850 posluchačů československé státní příslušnosti, kteří v roce 1965/6 navštěvovali 1. ročník řádného studia. (Výzkum byl realizován u 2726 studentů, tj. 95,7%.)

Při volbě fakult hrály svou roli i zkušenosti

Struktura výběrového souboru

Fakulta	Praha	Bratislava	Olomouc	Plzeň	Košice Prešov	+Součet
Strojní	150	150	—	140	168	608
Politické ekonomie	150	150	—	—	—	300
Přírodovědecká	150	150	150	—	144	594
Lékařská	150	150	150	150	150	750
Pedagogická	150	—	150	150	150	600
Součet	750	600	450	440	610	2.850

nosti z jiných prací o vysokoškolských studentech (jejich seznam je uveden v bibliografii projektu výzkumu), a dále i subjektivní názory členů pracovního týmu.

Málokdy bývá v sociologickém šetření vhodné použít záměrného výběru. Nebezpečí spočívá hlavně v tom, že při interpretaci se často zapomíná na omezení, která s sebou záměrný výběr nese, a již jen tento moment je potenciálním nebezpečím nesprávné interpretace empirických dat a formulace nesprávných, nepodstatných závěrů. Při interpretaci výsledků záměrného výběru lze generalizovat, avšak pouze ve výjimečných případech. I v tomto výzkumu lze zobecnit některá fakta pro celé studentstvo, avšak mohou to být úvahy jen některých typů (např. srovnávací) a je třeba při tom zachovávat značnou opatrnost.

Cíle výzkumu

Smyslem a cílem naší aktivity bylo nalézt ty parametry struktury činnosti studentů, které výrazně diferencují jednotlivé fakulty navzájem, ale které jsou společné pro určitá seskupení fakult, a též nalézt znaky, které nediferencují fakulty vůbec, nebo naprosto nevýznamné. Šlo tedy, mimo jiné, o zjištění, ve kterých ze zjišťovaných znaků se studenti devatenácti záměrně vybraných fakult sobě nejvíce „podobají“.

Z aspektu merita výzkumu jsme předpokládali, že celý soubor se bude projevovat jako nehomogenní, tj. že rozložení studentů na fakultách podle konkrétních znaků budou značně rozdílná.

Položili jsme si tedy zároveň novou otázku, která vlastně vyplývá z původně stanovených hypotéz (a), (b) (viz výše): zda je tu nějaký homogenizační činitel, zda lze celou původní skupinu fakult, u níž předpokládáme heterogennost, rozložit na podskupiny, které by již byly homogenní. Přitom nás ovšem nezajímaly libovolné skupiny, ale takové, které odpovídaly naší pracovní hypotéze a které se daly rozumně vysvětlit.

Vytvořili jsme podskupiny podle a) *lokalit* [fakulty umístěné v Plzni, Praze, Olomouci, Bratislavě, Košicích + Prešově], b) *oborů studia* [fakulta strojní, přírodovědecká, lékařská, pedagogická a fakulta politické ekonomie].

Výchozí hypotéza spočívala v tom, že obor studia by měl být relativně silným homogenizačním faktorem, neboť lze předpokládat, že již sama volba stejného typu fakulty je předpokladem určitých shodných rysů struktury osobnosti, jež jsou společným studiem neustále upevňovány. Současné zkoumání podsouborů podle lokalit bylo spíše doprovodným jevem procedury zkoumající první hypotézu. Mělo ve většině případů první hypotézu nepřímo podpořit, v některých konkrétních případech pak upozornit na faktory, které jsou homogenizovány seskupením podle lokalitního ukazatele. (Technická stránka této části spojení s druhou hypotézou proběhla se zanedbatelnými finančními náklady.)

Z určitého hlediska nás samozřejmě zají-

mal i samotný fakt frekvence výskytu v jednotlivých kategoriích zkoumaných znaků, a to četnosti jak v jednotlivých souborech samotných, tak ve skupinách, které vznikaly procesem homogenizace. O problémech s tím spojených se však zde nebudeme šířit.

Teprve na základě zjištění těchto typů mohlo být uvažováno o případných obecnějších tvrzeních o studentstvu. Z hlediska metodického jsme pohlíželi na jednotlivé fakulty jako na devatenáct základních souborů, v nichž je provedeno devatenáct vzájemně nezávislých náhodných výběrů. Každý náhodný výběr tedy vypovídal o určité fakultě a prostřednictvím těchto výběrových souborů jsme mohli provést srovnání příslušných devatenácti základních souborů.

Statistická procedura

Řešení sociologického problému popsaného v druhé části bylo provedeno za pomoci procedury testování statistických hypotéz. Jelikož popsaná úloha se zdá být velmi důležitá i v přenesení do jiných společenskovedních výzkumů, neboť umožňuje srovnávání souborů a jejich kategorizaci, a jelikož její použití není příliš rozšířené, věnujeme tu několik řádek popisu její matematicko-statistické paralely. K tomu účelu zde zavedeme některé pojmy, které korespondují se sociologickými pojmy používanými v první části.

Skupinu souborů (dvou či více) považujeme za statisticky homogenní, jestliže můžeme předpokládat, že statistická rozložení zkoumaného znaku jsou shodná, tj. jsou stejného typu a jsou určena stejnými parametry. V našem případě jsou soubory fakultami a zkoumaný znak je kvalitativní, který může nabývat dvou nebo více možných hodnot. Statistické rozložení ve třídách (kategoriích, hodnotách) můžeme tedy považovat za multinomické (pro dostatečně velké základní soubory) a jeho parametry vyjadřují pravděpodobnosti (resp. relativní četnosti) výskytu znaků v jednotlivých třídách. Homogenita skupiny souboru znamená tedy, jinými slovy, že percentuální zastoupení jednotek souborů v jednotlivých třídách je stejné nebo málo odlišné. To dále znamená, že soubory považujeme za statisticky homogenní, jestliže můžeme předpokládat, že jsou všechny vytaženy ze stejné populace. Jestliže soubory jsou určeny nějakým identifikačním ukazatelem (např. v našem případě určitým typem fakulty) a jestliže statistická rozložení v nich jsou shodná, pak máme důvod se domnívat, že společná vlastnost všech souborů, která určuje seskupení, je homogenizačním činitelem, a jednotlivé soubory ve skupině se nemusí dále dělit na skupiny menší. To umožňuje také rozumnou interpretaci marginálního rozložení.

Naproti tomu považujeme skupinu souborů za statisticky heterogenní, jestliže některý z nich má rozložení s jinými parametry než ostatní. V tomto případě si lze klást další otázku, který soubor či která podskupina homogenitu narušuje a musí být z celé skupiny vyčleněna.

Předložené úvahy vedly u většiny sledo-

vaných hledisek ke stanovení těchto simulánních hypotéz:

1. Celý soubor devatenácti fakult je nehomogenní.

2. Skupiny fakult stejného typu oboru studia jsou homogenní.

3. Seskupení oborů studia základních celků je nehomogenní.

Jako další homogenizační činitel bylo sledováno lokální seskupení.

Verifikace či odmítnutí těchto hypotéz byly provedeny na základě dat získaných statistickým testem, jenž byl odvozen na základě teorie informace (Kullback 1959).

Tento test umožňuje současná prověření všech hypotéz uvedených pod body (1), (2), (3). [Bod (2) obsahuje v sobě tolik hypotéz, kolik je utvořeno skupin.] Statistiku pro (1) lze rozložit na součet nezávislých statistik pro (2) a (3); tyto jsou ovšem zajímavé jen za předpokladu (1), který však byl ověřen ve všech případech.

Vzhledem k tomu, že tvar těchto statistik lze získat jednoduchým zobecněním jednoho případu z citované Kullbackovy knihy, neuvádíme jej zde. Tam jsou též odvozeny matematické základy metody, jejich teoretické rozložení a základní vlastnosti. Jelikož asymptotické rozložení je tu χ^2 -kvadrát, postačí pro testovací postup našich hypotéz běžné tabulky tohoto rozložení. Výhoda této procedury je též v tom, že výpočet hodnot testovaných kritérií je jednoduchý; jsou to sumy výrazů tvaru $n \ln$, jež lze získat z již vytvořených tabulek.

Tyto informační testy však umožňují stavět hypotézy nejen popsáním způsobem, jak jsme je stanovili v našem šetření vzhledem ke zkoumanému problému, ale i jinak podle nejrůznějších jiných potřeb. Tyto testy jsou však založeny na asymptotických vlastnostech, podobně jako χ^2 -kvadrát test, a proto musí být dodržena i stejná omezení na jejich použitelnost. To znamená, že musí být zajištěn dostatečný počet měření (dostatečné velikosti výběrových souborů) a ne příliš malé četnosti v jednotlivých polích fakulty.

K interpretaci lze obecně říci, že jestliže se nám ukáže některá skupina fakult jako homogenní z hlediska zkoumaného znaku, pak se neprojevují žádné rozdíly mezi soubory ve skupině a proto tu vystupuje zřejmě nějaký homogenizační činitel, jak jsme jej nazvali výše. Předpokládáme-li totiž, že kdyby skutečné rozdíly existovaly, pak by se projevil i statisticky. Jestliže se naopak projevují statisticky významné odchylky od homogenity, pak tu máme naznačeno, že společná vlastnost souboru skupiny jako homogenizační činitel nepůsobí, že je tu něco, co způsobuje nezanedbatelnou odchylku. Tento vliv již nelze vysvětlit náhodným působením. Příčiny mohou být různé — úvahy kauzálního charakteru jsou však dosti složité a nelze je dělat obecně; každý případ je třeba řešit zvlášť.

Některé výsledky a jejich interpretace

Vzhledem k rozsáhlosti dat, která byla k dispozici, není možné uvádět detailní závěry,

kteří budou rozpracovány v závěrečné zprávě výzkumu.

Omezíme se proto jen na sdělení některých výsledků. Typ fakulty se projevilo jako mnohem silnější homogenizační činitel než lokalita. Tento závěr vyplývá nejen z prosté frekvence výskytu „faktu homogenity“, ale především z frekvence jeho výskytu v elementech prioritní významnosti.

Z aspektu meritu výzkumu se jako relativně nehomogennější projevilo výběrový soubor v oblasti politických postojů. V této meritorní oblasti byla zjištěna vysoká frekvence shody struktur kvalitativních znaků. Naopak oblast studia, studijní aktivity, jejího charakteru, i studijní výkon byly oblastí výrazné heterogenity. V této oblasti se projevil výrazné variace nejen v rámci celého výběrového souboru, ale i v jednotlivých podsouborech fakult a lokalit.

Z tohoto zjištění vyplývá, že z daného aspektu byly základními diferencujícími znaky výběrového souboru elementy studijní činnosti, v nichž se výrazně projevilo specifikum jednotlivých konkrétních fakult (rozdílnost vstupního materiálu, rozdílnost charakteru práce vysokoškolských pracovníků, hodnocení výkonu studentů, variace v charakteru studijní aktivity posluchačů atd.).

Z metodologického aspektu vedly výsledky testování homogenity k zpřesnění požadavků na další zpracování empirických dat (třídění vyššího stupně, výpočet 2. norm. koef. kontingence), k určení, v kterých elementech je možno pracovat s výběrovým souborem jako s totalitou, a které oblasti vyžadují diferencovaný přístup.

Chtěli bychom tu poznamenat, že výsledky této procedury naznačily možnosti globalizujících přístupů k studentstvu, ovšem jen z určitých hledisek, a též možnosti velmi opatrné generalizace některých faktů pro širší soubor celého studentstva.

Závěrem však ještě malou poznámku k typu takového šetření. Způsob uspořádání testu a jeho interpretace naznačují, že využití tohoto postupu je daleko širší než jen pro výzkum vysokoškolských studentů. Užitečnost této cesty se objevují všude tam, kde do výzkumu přistupují dva (případně více) relevantní faktory, které je třeba souběžně zkoumat. Zde jsou jimi lokalita a instituce, a zkoumáním homogenity pak vytváříme důležité homogenní typy uvnitř obou hledisek. V jiných problémech to mohou být faktory jiné. Sledování populace z hlediska lokálního a institucionálního (poznamenejme, že slova lokalita a instituce tu můžeme chápat velmi široce) však je užitečná téměř vždy, když zkoumáme statistickou strukturu, neboť tento postup nám umožňuje odhalovat kontexty spojené s oběma faktory, a to současně jedním stupněm ekologické analýzy zkoumané populace.

Literatura:

S. Kullback, *Information Theory and Statistics*, Wiley & Sons, New York 1959.

Eliška Freiová, Jan Řehák