

Oblastní výběrová šetření v sociologii

J. P. Voronov

Novosibirsk

Vědecké plánování a řízení společnosti se neobejde bez podrobných informací, jejichž rozsah je podstatně větší než údaje uváděné v pravidelných státních hospodářských zprávách.

Při podrobnějších výzkumech života společnosti jsou nutné jednorázové průzkumy, které poskytují materiál pro ekonomická rozhodnutí. Všeobecně se uznává, že taková *jednorázová šetření* musí být *výběrová*, protože výběr podstatně snižuje náklady na *výzkum* a někdy i zvyšuje jeho přesnost.

Sociální jevy vyžadují operativní rychlé zachycení mnohých parametrů a to je při celkové evidenci obtížné, často dokonce nemožné. Proto je vědecký výzkum metodických problémů konstrukce výběru v teoretickém i praktickém smyslu důležitý. Získali jsme mnoho zkušeností jak v oblasti praktických aplikací, tak i ve zpracování metodických problémů *uspořádání výběru*.

Sovětské statistické, kteří pokračují v bohatých tradicích zemské statistiky, dosáhli v rozvoji výběrové metody významných úspěchů.

Zkušenosti získané využitím výběrů při zkoumání sociální sféry umožňují stanovit nejrozsáhlejší schémata výběru.

Z celého souhrnu metodických problémů výběrových šetření je v referátu dotčena jen část problému uspořádání oblastních výběrů. Je nutno zdůraznit, že právě zpracování tohoto okruhu problémů je neobvykle důležité.

Většina výběrových šetření poskytujících materiál pro rozhodování v oblasti ekonomie a sociologie má specifickou a jednotnou podobu. Výběry v takových šetřeních jsou obvykle jednorázové a vícestupňové, přičemž na nižších stupních se používá náhodného výběru, na vyšších stupních výběru oblastního nebo výběru typických prvků. Tyto zvláštnosti konstrukce výběru v demografické statistice platí i pro zjišťo-

vání sociálních problémů práce a pracovních zdrojů a po dlouhá léta zůstávaly téměř nezměněny. Je nutno je považovat za správné, neboť se vytvořily v důsledku zdokonalování různých schémat výběru v aplikaci na konkrétní šetření. Postupné zdokonalování vedlo k dílčím zlepšením v rámci analyzovaných principů výběru.

V současné době je nejaktuálnější zdokonalování metodiky výběru na vyšších stupních, jednak proto, že v poválečných letech byl vypracován nový matematický aparát, který umožňuje provádět formalizaci výběrů na vyšších stupních, jednak proto, že rozšíření sféry aplikace metody výběru nutně vyžaduje — pokud jde o sociální statistiku — nové formy na vyšších stupních.

Zdokonalení, kterých bylo v poslední době dosaženo, je možno shrnout do dvou skupin:

1. využití matematických metod výběru typických objektů,
2. optimální stratifikace.

Tyto dvě skupiny se navzájem liší, ačkoli řeší stejný úkol — zapojení formálního aparátu do té etapy práce, v níž se dosud fakticky nepoužíval.

Na základě analýzy historického materiálu je možno se zmínit o těchto vývojových tendencích u uspořádání výběrových šetření, podstatných pro další zdokonalení metodiky:

1. Princip spojování oblastního výběru nebo výběru typických prvků (ve vícestupňovém výběrovém šetření) na vyšších stupních s náhodným výběrem na nižších stupních se vytvořil již v zemské statistice a aktivně ho bylo použito v prvních sovětských výběrových šetřeních.

2. V prvních sovětských šetřeních byl stanoven princip předběžného určování výběru proměnných, podle něhož se provádí stratifikace a typizace.

3. Zatímco použití náhodného výběru na

nižších stupních je způsobem obecně uznávaným, o principech výběru na vyšších stupních se často diskutuje a znovu se prověřují.

4. V posledních letech se znatelně rozšiřuje sféra aplikace oblastních výběrů a zvyšuje se jejich podíl.

Z uvedených tendencí jasně vyplývá, že se analýzy provádějí pomocí nejrozšířenějšího a nejperspektivnějšího typu výběrů a že zvláštní pozornost je nutno věnovat výběru na vyšších stupních, protože právě tyto metody výběru se méně stabilizovaly.

Při konstrukci výběrů v sociologických výzkumech se používá velmi různorodých přístupů, což je způsobeno především tím, že samy sociologické výzkumy jsou teprve v počátečním stadiu.

Konkrétní sociologické výzkumy jsou stejně jako výzkumy státní statistiky víceúčelové a nepřinášejí zásadně nové metody konstrukce výběru. Používají v současné době stejných metod výběru jako sovětská státní statistika. Sociologové by tedy měli více využívat zkušeností statistických orgánů.

Zároveň však je možno ukázat na příkladu konstrukce výběrů při výzkumu věnovanému vztahu pracující mládeže k práci [1], že v konkrétních sociologických výzkumech se rodí nový princip pro uspořádání výběru. Tento princip lze formulovat takto: proměnnými stratifikace nebo typizace musí být znaky odpovídající pojmům, jež jsou součástí základní hypotézy výzkumu. Ani v jedné sociologické práci se tento princip přímo neuvádí, ale zmíněná práce leningradských sociologů má k tomu blízko. Nepřesnosti u dat výběrového souboru jsou často způsobeny nedůsledností při realizaci tohoto principu.

Zadání stratifikačních znaků ve shodě se základní hypotézou šetření se může chápat jako proces zpřesňování úkolů mnohoúčelového výběru, jasně určení okruhu zkoumaných problémů.

Existují dvě základní reprezentace apriorní informace: a) ve tvaru vzdálenosti (měr blízkosti) mezi objekty a b) ve formě znaku.

Problém přechodu od jedné formy reprezentace ke druhé ještě není vyřešen a úkoly oblastního výběru objektů se řeší různými metodami v souvislosti s formou reprezentace apriorní informace.

Rovněž není ještě jasně formulován problém unifikace apriorní informace před uspořádáním výběru. Zdá se, že pojem prostoru charakteristik, na který poukázal P. Lazarsfeld a G. Barton [2]), se může stát dobrým základem pro standardní reprezentace apriorní informace. Rozvinutím tohoto pojmu jsou pojmy „konstruktivní“ a „operativní“ prostor charakteristik, důležité pro další práci [3].

Konstruktivní prostor charakteristik je unifikovaná forma apriorní informace, na jejímž základě se provádí výběr objektů. *Operativní prostor charakteristik* je systematizovaný zápis těch znaků, jejichž hodnoty jsou pevné v průběhu šetření [2]. V souvislosti se způsobem přechodu od konstruktivního prostoru k operativnímu můžeme vydělit dva zásadně různé typy konstrukce výběru.

Jeden z nich můžeme nazvat výběrem typických prvků. Z celého souboru objektů se vybírají objekty typické a ty se pak podrobují všestrannému zkoumání. Druhým typem je oblastní výběr, při němž se sjednocují objekty do oblasti a určují se výběrové poměry uvnitř oblastí. Tyto dva typy výběru dobře rozlišíme za pomoci pojmů „operativního“ a „konstruktivního“ prostoru charakteristik. Jsou-li v konstruktivním a operativním prostoru charakteristik určeny stejné objekty, pak práce na uspořádání výběru prakticky znamená přechod od jednoho zkoumání ke druhému, značí splnění úkolu, který se týkal zvětšení rozsahu informace a objektu. Jsou-li v konstruktivním prostoru objekty jiné než v prostoru operativním, práce na konstrukci výběru se realizuje jako výběr určité části stupňů výběru při vícestupňovém uspořádání.

Všimněme si úkolu vytypování skupin, které jsou nejtěsněji spojeny s ostatními skupinami při existenci měr blízkosti mezi všemi skupinami.

Úkol může být vyřešen metodou teorie grafů [4]. Zde však vznikají nové, složitější problémy. V praktických příkladech byly počítány zpočátku míry blízkosti (míry podobnosti) mezi skupinami a pak se podle hodnot těchto měr stanovily dvojice skupin, které se navzájem podobaly, a na grafu, jenž takto vznikl, se vydělovaly tyto skupiny tak, že každá ze zbylých skupin se podobala alespoň jedné skupině z tohoto (a pouze z tohoto) podvýběru.

Takový způsob je však složitý, neboť není známa hodnota míry blízkosti, při níž je možno považovat dva prvky za vzájemně podobné. Tento problém není ještě ukázán v plném svém rozsahu. Nicméně však jsou načrtnuty základní způsoby jeho řešení, které v podstatě znamenají rozložit množinu měr blízkosti na dvě třídy ve shodě s optimálním výběrem.

Přístup ke stratifikaci je možný z hlediska metod rozpoznávání obrazů, které se v poslední době rozvíjejí. Nejdříve je třeba formulovat úkol stratifikace ve shodě s používanými algoritmy rozpoznávání obrazů a pak přistoupit k řešení tohoto úkolu. Jako příklad můžeme uvést konstrukci výběru pro výzkum příčin migrace obyvatelstva z venkova do měst, který provedl *Institut ekonomie a organizace průmyslové výroby Sibiřského oddělení Akademie věd SSSR společně s Ústředním statistickým úřadem RSFSR v Novosibiřském kraji* v r. 1961.

Výběr byl třístupňový. První stupeň znamenal výběr okresů, druhý stupeň byl výběrem vesnických sovětů a třetí stupeň výběrem rodin.

Na třetím stupni výběru byl využit systematický výběr z hospodářských knih. Na prvních dvou výběrových stupních se pro stratifikaci používalo metody taxonomie. Rozhodnutí o výběru znaků pro konstruktivní prostor charakteristik bylo provedeno na základě diskuse, již se zúčastnili jak vědečtí pracovníci, tak i pracovníci CSÚ.

Nakonec bylo pro výběr administrativních oblastí určeno pět znaků (konstruktivní prostor I):

1. početnost vesnického obyvatelstva v okrese na počátku roku šetření,
2. procento kolchozních usedlostí z celkového počtu rolnických usedlostí,
3. relativní saldo migrace venkovského obyvatelstva za posledních šest let před šetřením,
4. vzdálenost okresních center od center kraje,
5. vzdálenost okresního centra od nejbližší železniční stanice.

Pro výběr vesnických sovětů bylo určeno šest znaků (konstruktivní prostor II), které se shodují se znaky určenými pro okresy.

Na každém stupni se stanovily orien-

tační části výběru a pak v konstruktivních prostorech (zpočátku v prvním, pak ve druhém) se množiny objektů dělily na třídy. Počet tříd se určoval počtem jednotek, které bylo nutno vybrat podle orientačních částí výběru na příslušných stupních, protože z každé takové třídy („taxonu“) se vybírala jedna jednotka, která byla pro ni typická. Tento postup byl předmětem nejedné diskuse.

Uvedeme si autorovy argumentace v těchto diskusích:

Existují dvě možné zásady výběru, které se při nerovnoměrném rozložení celkového souhrnu v konstruktivním prostoru charakteristik navzájem vylučují.

První princip spočívá v maximálně stejné možnosti zahrnutí každého prvku do výběru.

Druhý princip představuje maximálně stejnou možnost zahrnutí prvků pro každou oblast konstruktivního prostoru charakteristik. Zavedení proměnné části výběru podle taxonů — tříd (proporcionálně jejich velikosti) znamená realizaci prvního principu. Výběr podle jednoho typického objektu znamená realizaci druhého principu.

Protože se využívá materiálů sociální statistiky hlavně pro strukturní srovnávání, při nichž je důležité zachování různých podmínek existence objektů, pokládá se za důležitější druhý princip, a je tudíž třeba vybírat po jednom typickém objektu z třídy (taxonu). Při srovnávání výběrových a všeobecných šetření podle základních charakteristik obyvatelstva (demografických a jiných) se mezi nimi prokázala těsná shoda.

Druhý princip — ve světové praxi rozšířenější — je zdokonalování výběru na vyšších stupních. Tento způsob je zaveden v zahraniční statistice a je možno poukázat na dvě hlavní vědecké školy, které ho zpracovávají. První z nich — skandinávská škola matematické statistiky — vypracovává formální schémata optimální stratifikace, druhá — indická škola statistiků-praktiků — má velké zkušenosti s výběrovým šetřením při různých formách stratifikace.

Rozdíl v principech konstruování jedno- a mnohorozměrných výběrů je patrný, a proto přechod od jednorozměrných výběrů k mnohorozměrným je spojen s obtížemi. Bezprostřední úkol nalezení optimálního oblastního výběru zna-

mená určit hranice oblastí — při daném rozsahu výběru a počtu oblastí — a rozmístit výběr v oblastech tak, aby účelová funkce dosáhla svého optima.

Analogicky se formuluje i duální úloha: pro přijatelnou hodnotu účelové funkce a daný počet oblastí dosáhnout takového výběru hranic oblastí a rozmístění výběru podle oblastí, které by minimalizovaly rozsah výběru.

Kritérium optimálnosti, které je v zahraniční statistické praxi nejrozšířenější a je založeno na minimalizaci výběrového rozptylu, poprvé použil v r. 1950 švédský statistik T. Dalenius [7]. Od té doby se toto kritérium nejednou modifikovalo. Velká část těchto modifikací souvisí se zjednodušením výpočtů. Řešení úkolu optimálního výběru podle návrhu Dalenia se neobejde bez složité výpočetní techniky. Zjednodušení, jakých se při konkrétní práci používá, neumožňují sice dosáhnout přesné optimálního řešení (jak se prokázalo v celé řadě prací a bylo prověřeno praktickými přehledy [8] [9]), avšak dosahované aproximace se blíží optimálnímu řešení.

Autor prováděl optimální oblastní výběr založený na minimalizaci rozptylu pro tři rozsáhlé průzkumy (podrobně viz v [9]). Jako první je možno uvést průzkum čtenářů časopisu *Pravda*, kde probíhala optimální stratifikace v rámci tří svazových republik: Ruské socialistické federativní sovětské republiky, Ukrajiny a Kazachstánu. Jako proměnná oblastního výběru byla zvolena hustota subskripce novin. Tento znak se považoval za vhodný pro úkol, který měla na zřeteli redakce novin. Podobný byl i druhý úkol: konstrukce optimálního oblastního výběru pro čtenáře časopisu *Litěraturnaja gazeta*. Oblastní výběr se také prováděl podle znaku „hustoty předplatného novin“, dělil se však na oblasti současně celý komplex svazových republik a také území RSFSR, Ukrajiny a Kazachstánu. Byl učiněn pokus konfrontovat oblastní výběr podle hustoty předplatného s ekonomickým oblastním výběrem. Největší shoda mezi oběma výměry byla zaznamenána při oblastním výběru podle hustoty předplatného novin *Pravda* v oblasti Ukrajiny. Třetí úkol se týkal průzkumu mladých vědeckých pracovníků Sibiřského oddělení Akademie věd SSSR. Výběr se prováděl v institutech Sibiřského

oddělení Akademie věd podle dvou znaků charakterizujících „věkovou příslušnost“ pracovníků: střední věk vědeckých pracovníků a podíl kandidátů mladších 29 let z celkového počtu kandidátů.

Rozdělením ústavů SO AN do čtyř skupin podle těchto znaků se vydělily skupiny ústavů dostatečně stejnorodé i co do vědeckého zaměření.

Problémy *mnohorozměrných* optimálních oblastních výběrových šetření pomáhají řešit metody lineárního a kvadratického programování. (Byl vybrán Rosenův algoritmus kvadratického programování — s některými modifikacemi [10]).

Nakonec uvedeme stručně hlavní závěry, které vyplynuly ze zkoumání problémů, jimž jsme věnovali pozornost v tomto referátu.

1. Vícetupňové oblastní výběry jsou jedním z nejperspektivnějších schémat získávání informací o sociálních jevech.

2. Lze očekávat, že oba způsoby zdokonalování metodiky výběru na vyšších stupních (vyčleňování typických skupin a optimální oblastní výběr) přinesou žádané výsledky.

Z ruštiny přeložila Milena Manová

LITERATURA

- [1] *Čelovek i jeho rabota*. Pod redakcí Zdravomyslova A. G., Rožina V. N., Jadova V. A., M., 1967
- [2] *Language of Social Research*. Ed. by P. Lazarsfeld and G. Barton, New York, 1956, s. 40-53
- [3] Četverikov N. S., *Statističeskije i stochastičeskije issledovanija*, M., 1963
- [4] Ore, O., *Theory of Graphs*, Providence, 1962
- [5] *Metodika vyboročnogo obsledovanija migracii selskogo naselenija*. Pod red. T. I. Zaslavské, V. D. Mirkina, K. F. Jeršovové, Novosibirsk, 1969
- [6] *Raspoznavanije obrazov v socialnych issledovanijach*. Pod red. T. I. Zaslavské, I. G. Zagorujka, Novosibirsk, 1968
- [7] Dalenius T., *The Problem of Optimum Stratification*, Skandinavsk Aktuarie tidsskrift, 1950, čís. 3-4 s. 203—213
- [8] *Doklady vsesozuznomu simpoziumu po sociologičeskim problemam sela*. Pod red. T. I. Zaslavské, Novosibirsk 1968
- [9] Gren I., *O pewnym rastosowaniu programowania nieliniowego metody reprezentacyjnej*. Przegląd statystyczny, 1966, t. 13, n. 3
- [10] Rosen I. B. *Pattern Separation by Convex Programming*. — Journal of Mathematical Analysis and Applications, 1965, v. 10 N. 11 s. 123-134