

Kvótní výběr patří mezi metodologické postupy, na něž neexistuje jednoznačný názor. Je to metoda, která má své velmi silné odpůrce, ale současně i své jednoznačné a jednostranné příznivce. Historie metody je poznamenána mnoha prokazatelnými úspěchy, ale i mnoha zcela jasnými a naprostými zklamáními. Její vývoj je spjat s vývojem metodologie empirického sběru informace pro potřeby společenskovědního výzkumu, výzkumu veřejného mínění, výzkumu trhu apod. Počátky tohoto vývoje lze hledat kolem roku 1936 a lze je spojit se jmény Cherington, Roper, Gallup a Crossley, jejichž výzkumy se staly známými především po úspěšné predikci výsledků amerických prezidentských voleb. Od této doby se metoda značně vyvinula a zároveň, používáním v různých podmínkách, i velmi diferencovala. Dnes je možné charakterizovat kvótní výběr spíše jako celý komplex metod a postupů než jako jedinou metodu.

Kvótní výběr se velice rychle rozšířil do nejrůznějších aplikací. Důvodem pro to byly především relativně nízké náklady a jednoduchá organizace a popularita přesných a praxí ověřených výsledků. Přesto mezi metodology, matematickými statistiky a profesionálními výběraři vždy převládaly a stále převládají nedůvěra k tomuto postupu a značně silné metodologicko-statistické výhrady. Praktické aspekty, organizace, náklady a další podmínky nás však nutí, abychom kvótní výběr používali. Uživatelé metodu vychvalují, matematici ji kritizují. Účelem této stati je metodu charakterizovat a dát každému uživateli možnost rozhodnout se na základě vlastního studia problému.

Shrnující názor na kvótní výběr vyjadřuje L. Kish (1965:563) takto: „Kvótní výběr není jednoznačně definovaná vědecká metoda. Je to spíše umělecký výtvar, stěží definovatelný či popsateľný. Obecná kritika nemůže jít proto do detailů. A specifická kritika jednoho postupu se nemusí hodit na jiný postup“. Toto je bezesporný fakt. A snad právě zde se skrývá i důvod, proč znalosti o kvótním výběru a jeho možnostech, přednostech i nedostacích se tolik rozcházejí, proč jsou u většiny uživatelů nepřesné a často i chybné.

Účelem této stati je metodu popsat v jejich obecných principech, charakterizovat ji vzhledem k celkové reprezentativitě zadané výzkumnými cíli a formulovat problémy, které je nutno v praxi vyřešit, máme-li s klidným svědomím zobecňovat data získaná kvótním výběrem. Půjde o pokus odpovědět na otázky:

- 1) Co je to kvótní výběr?
- 2) Jaké jsou základní etapy v postupu kvótního výběru?
- 3) Co podmiňuje zobecnitelnost kvótního výběru?
- 4) Jaké rušivé vlivy působí na různých místech postupu?

1. Charakteristika kvótního výběru

Při plánování kvótního výběru jsme vedeni myšlenkou zajistit „téměř“ úplnou shodu výběrového a základního souboru vzhledem k vybraným znakům a jejich

populačním rozložením. Chceme zajistit shodu procentního rozložení několika základních sociodemografických znaků v populaci a ve výběru tím, že tazatelům předepíšeme kvóty (tj. počty osob v kategoriích daných znaků), které při výběru musí dodržet a které jsou voleny tak, aby požadovanou shodu zajistily.

Pro zajištění reprezentativity však shoda u vybraných, tzv. kvótních znaků nestačí. *Výběrový postup je aplikován za předpokladu, že shoda u vybraných znaků zajistí i reprezentativitu u ostatních znaků.* Na platnosti tohoto předpokladu závisí použitelnost metody, a právě kolem tohoto předpokladu se odvíjí proces rozhodování o přijetí a použití postupu.

Tento předpoklad nelze obecně ani jednoznačně přijmout ani jej nelze jednoznačně zamítnout. V praxi nikdy úplně neplatí, chceme se však k němu co nejvíce přiblížit. A právě zde je prostor k uplatnění onoho umění organizátorského i statistického, umění využít všech dostupných statistických informací, teoretických znalostí a zobecnění předchozích výsledků a zkušeností, psychologicko-výchovného umění při práci s tazateli a vůbec uplatnění celkové sociologicko-statistické erudice.

Pro hodnocení reprezentativity [Řehák 1978] výběrového šetření i pro rozhodnutí o způsobu analýzy a zobecnění výsledků je nutno znát matematický model procesu sběru dat, aby bylo možno odvodit jeho důsledky a použít je při výpočtech odhadů statistických ukazatelů i určení nepřesností těchto odhadů. U kvótního výběru se setkáváme s principiální potíží, spočívající v tom, že není znám vhodný matematický model, který by proces vzniku dat charakterizoval a z něhož by tudíž bylo možno odvodit i reprezentativitu. Zatímco pravděpodobnostní výběr je procesem řízené náhody, kvótní výběr by měl srovnatelný model jen v tom případě, kdyby tazatelé fungovali v rámci svých kvót, předpisů a umístění jako perfektní generátory náhodného výběru. Takovou vlastnost ovšem tazatelé nemají a ani mít nemohou.

Kvótní výběr bývá někdy srovnáván se stratifikovaným (oblastním) pravděpodobnostním výběrem. Kvótní znaky mají skutečně do značné míry charakter stratifikačních proměnných⁽¹⁾ a také z této vlastnosti při jejich výběru a určení vycházíme. Každý kvótní znak či kombinace znaků mají (jednotlivě) stratifikační význam. Souhrn všech předpisů může však mít efekty daleko složitější. Kvótní výběr však jako stratifikovaný charakterizovat nelze, neboť uvnitř oblastí nevybíráme respondenty náhodně. Vzhledem k těmto aspektům nemůžeme ani přesně určit *stratifikační efekt*, který použitím kvótních znaků získáváme. Přesto předpokládáme, že stratifikační efekt existuje a že působí na zvýšení přesnosti výsledných odhadů, a chceme jej maximalizovat.

Proti stratifikačnímu efektu působí *efekt skupinkový*, který vzniká u jednotlivých tazatelů. Tazatel, je-li mu dopřána volnost, vybírá osoby s nimiž snadno navazuje kontakt, z určitého prostředí, respondenty, které snadno vyhledá a u nichž nejsnadněji identifikuje charakteristiky kvótních znaků. Tazatel tak,

(1) Stratifikací rozumíme ve statistice rozklad populace do několika oblastí (skupin statistických jednotek tvořících populaci), které jsou z hlediska daného problému vnitřně co nejhomogennější a navzájem se naopak co nejvíce liší. Proměnná, jejíž hodnoty určují jednotlivé oblasti, se nazývá stratifikační proměnnou. Oblasti mohou být určeny i pomocí kombinace několika proměnných. Formálně mohou být určeny libovolným rozkladem populace, praktický význam mají však jen ty rozklady, které mají výše uvedenou vlastnost. Proto za stratifikační proměnné volíme takové znaky, které jsou co nejsilněji korelovány s danými problémy (co nejsilněji diferencují vzhledem k danému problému). Je-li stratifikační znak vhodně volen (má silnou diferenciační roli), výsledné odhady mívají podstatně vyšší přesnost.

být i neuvědomněle, je nositelem skupinkového efektu, tj. homogenizuje svůj malý soubor dotázaných. Hlediska homogenizace přitom nejsou kontrolovatelná, stojí mimo kvótní znaky. Oba tyto efekty, stratifikační a skupinkový, působí z hlediska přesnosti závěrů proti sobě. Zatímco stratifikační efekt zvyšuje přesnost jednotlivých odhadů, skupinkový efekt působí naopak. (2)

Oba aspekty výběru je však nutno hodnotit v souvislosti s třetím, velice vážným a nepříjemným aspektem, který se projevuje v *efektu vychýlení*. Tento efekt znamená systematické posunutí výsledků určitým směrem (Řehák 1978). U kvótního výběru je tento efekt o to nebezpečnější, že je velmi obtížně identifikovatelný. (3)

Na efekt vychýlení působí tyto faktory:

a) *dostupnost*,

tj. tendence tazatelů změnit cílovou populaci na tu její část, která je snadno dostupná.

Tazatelé :

— se obracejí na osoby, které najdou doma či na místě určeném pro rozhovor (jestliže nikoho nenajdou, na místo se již nevracejí);

— mají tendenci vybírat osoby z co nejbližšího okolí, aby minimalizovali časové ztráty;

— mají pro rozhovor (a pro výběr) k dispozici pouze určitou část dne (například večer).

Tento faktor má velmi silný vliv. Například budou-li tazatelé provádět rozhovory pouze večer, bude podstatně sníženo procento pracovníků, kteří chodí do práce na více směn (a to může být silně korelováno s tématem výzkumu); totéž se týká aktivistů a funkcionářů, kteří v období příprav velkých akcí (například spartakiády) nebyvají večer doma, čímž dochází ke zkreslení zastoupení těchto osob a tím i ke zkreslení zjišťovaného názoru populace.

b) *Ochota respondentů, odmítnutí*

Odmítnutí, které je korelováno s problematikou výzkumu či s postojem k výzkumu vůbec, je jednou z nejnepříjemnějších komplikací metody. Působí na vychýlení a zkreslení přímo, a to většinou tak, že lidé se záporným postojem odmítají výpověď, takže získáme nadhodnocené výsledky. Nepřímý vliv se pak projevuje i v mechanismu výběru respondentů: tazatelé podvědomě navazují kontakt s lidmi, u nichž předpokládají (na základě zkušenosti), že rozhovor proběhne hladce a že nebudou odmítnuti. To pochopitelně zesiluje efekt vychýlení. (4)

c) *Preference při výběru respondentů*

mají vliv na vychýlení směrem k lidem, kteří jsou tazateli sympatičtí, s nimiž

(2) Je to způsobeno tím, že vhodná stratifikace (tj. volba kvótních znaků) přispívá k plnému pokrytí variability dat (stratifikace nemůže zhoršit přesnost odhadů), zatímco homogenizace ve skupinách vede k relativní ztrátě informace o rozmanitosti a tím snižuje efektivnost výzkumu (snižuje se tak zvláště efektivní velikost výběru).

(3) Identifikovat jej můžeme vlastně pouze tehdy, máme-li konkrétní nevychýlené výsledky získány jinou nezávislou kontrolní metodou (buď nezávislým pravděpodobnostním šetřením, či z objektivních celostátních statistických údajů). Identifikaci vychýlení (a obecněji: validizaci celého postupu) lze provést tím, že výsledky interpretace a z ní odvozené prognostické závěry jsou konfrontovány s vývojem zkoumaného sociálního jevu. Takový postoj však je zvláště obtížný, neboť sem vstupují další faktory, které mohou neočekávaně vývoj jevu změnit.

(4) Je proto nutné, aby tazatelská síť byla neustále a plynule obměňována. U zkušených a dlouhou dobu pracujících sítí tazatelů se může projevit zdánlivě kladná okolnost: velmi nízké procento odmítnutí, které je však způsobeno právě (i neuvědomělou) preventivní snahou tazatelů o co nejjednodušší splnění úkolu.

se dobře spolupracuje; nejde jen o zkušenost s neochotnými typy lidí, ale o obecné preference každého tazatele.

d) *Poctivost a kvalita tazatelské sítě*

se projevuje v efektu vychýlení v různých formách:

- tazatel vyplňuje záznamové listy sám, bez respondentů (to vede k tomu, že tazatel s malými odchylkami zaznamená takové názory, o nichž předpokládá, že se vyskytují v populaci)
- tazatel vybírá osoby blízké (příbuzné, přátele, sousedy, spolupracovníky)
- tazatel vytváří malý panel (chodí stále ke stejným osobám)
- tazatel provede rozhovor, ale osobnostní charakteristiky upraví tak, aby odpovídaly kvótám.(5)

Tento faktor vychýlení je při pečlivé kontrole ve většině svých forem identifikovatelný. Nutno si však připustit, že příliš „dobré“ výsledky tazatelské sítě, jako je naprostá shoda kvótních požadavků, malý počet odmítnutí apod., mohou být nepřírozené — kvalita se pak projevuje pouze ve formálních ukazatelích, skutečné hodnoty získaných dat jsou problematické.

e) *Chybné odhadové postupy*

Odhadové postupy nemusí odpovídat postupu výběru, při nedodržení kvót či při nadnesení kvót u některých skupin obyvatelstva je nutné získaná data vážit. Volba nevhodné výpočetní formule (např. výpočet prostého procenta ze souhrnu dat) může vést k vychýleným výsledkům.

2. Proces kvótního výběru

Kvótní výběr je souhrnem opatření a postupných kroků, z nichž každý má svoji důležitost a každý může být zdrojem chyb. Jsou to metodologická, praktická rozhodnutí. Některá provádí výzkumník, některá tazatel:(6)

- A) určení kvótních znaků,
- B) určení proporcí pro jednotlivé kvótní znaky i jejich kombinace a určení kvót pro jednotlivé tazatele,
- C) výběr tazatelů,
- D) výběr respondentů,
- E) hodnocení reprezentativity,
- F) analýza dat,
- G) metodologické vyhodnocení jako modifikační činitel pro další kvótní výběr.

Úspěšnost kvótního výběru je podmíněna splněním mnoha nutných podmínek. Uvedeme si nyní metodologické zásady, které lze pro každý z uvedených kroků formulovat. Porušením každé zásady vnášíme do dat nekontrolovatelná zkreslení. Zásady jsou stanoveny s cílem působit na tři efekty:

(5) Sebelépe stanovené kvóty nepomohou, nebo se dokonce projeví zcela naopak jako silný faktor vychýlení, když tazatelé budou měnit osobní charakteristiky respondentů tak, aby jim vyšel kvótní předpis. Příčinou takového počínání může ovšem být také příliš komplikovaný kvótní předpis!

(6) Poznamenejme, že uvedené body se nevztahují k výzkumu jako celku, ale k jeho jedné vydělené části, kterou je kvótní výběr. Proces výzkumu je obsáhlejší. To je třeba mít na paměti, neboť v jednotlivých krocích hodnoceného postupu se ostatní aspekty výzkumného projektu velmi silně promítají: určující roli má především teoretický základ a cíl výzkumu.

- mají posílit efekt stratifikace
- mají zeslabit skupinkový efekt
- mají potlačit efekt vychýlení.

A) Určení kvótních znaků

Znaky pro určování kvót se vyskytují v předpise ve dvou možných formách:

- jednotlivé znaky vstupují do předpisu se svými jednorozměrnými populačními rozpořazeními
- znaky se v předpisech vyskytují v kombinacích (dvou nebo vícerozměrných)

U prvního postupu zajiřujeme shodu výběru a populace pouze z hlediska každého znaku zvlášť; předpis je relativně jednoduchý, ale nezajiřuje shodu dvourozměrových rozložení kvótních znaků. Je pochopitelné, že druhý přístup je metodologicky spolehlivější, je však prakticky náročnější. Rozhodnutí, který z obou přístupů zvolit, je však často závislé na počtu hodnot kvótních znaků: jde o to, zda zachovat podrobné členění u jednorozměrových kvót, či se spokojit s hrubším členěním (například věkových skupin) při možnosti volit kombinace.

Konečné slovo však závisí na konkrétních zkušenostech v daných sociálních, výzkumných, organizačních podmínkách i možnostech.

Zásady pro určování kvótních znaků

1. Kvótní znaky volíme tak, aby byly silně korelované s výzkumnými proměnnými.

Čím vyšší je korelovanost kvótních a zkoumaných znaků, tím vyšší je stratifikační efekt a tím silnější obranu proti vychýlení vytváříme. Dva aspekty této zásady plynou z definice znaku; jsou to:

- a) volba vlastností*, které budeme kontrolovat (věk, pohlaví, sociálně třídění, určení, geografické umístění, administrativní členění, velikost místa bydliště, obor národního hospodářství, ekonomická aktivita, vzdělání, studium apod.),
- b) volba hodnot znaku* (věkové skupiny, kategorie tříd a sociálních skupin, geografické či správní celky apod.).

Volba vlastností je primární, volba hodnot je však stejně podstatná. Nejjasněji je to vidět u věkových skupin: důležitost skupin i posun hranic se mění podle problémů a obsahových okruhů výzkumu.

Z hlediska stratifikačního efektu nemá význam zařazení kvótních znaků, které nekorelují s výzkumnými proměnnými (nekoreluje-li pohlaví se zkoumanými znaky, pak přes všechny zvyklosti v používání je přínos tohoto znaku k uvedeným efektům mizivý). I takové znaky však mají svůj význam: omezují libovůli tazatele a ruší skupinkový efekt (mají částečný vliv i na omezení efektu vychýlení). Jejich využití však není plně efektivní — k plné efektivitě je nutno volit znaky co nejsilněji korelované s tématem.

Z řečeného snad přímo plyne, že za určitých podmínek (především za předpokladu kumulace empirických zkušeností a jejich teoretických zobecnění) je možné k daným tématům (jednomu, více tématům současně či univerzálně) hledat soustavu kvótních podmínek optimálně tak, aby byl zaručen co nejvyšší stratifikační efekt. Hledáme přitom soustavu kvótních znaků, které představují optimální sociálně demografický korelát (resp. prediktor) sociálních jevů vcelku, či určitého výseku sociální reality odpovídající danému tématu výzkumu (pochopitelně při splnění ostatních zásad výběru kvótních

znaků.[7] Takováto činnost bude dnes asi spjata s vytvářením automatizovaných postupů určení kvót a výběru tazatelů počítačem.

2. Počet kvótních podmínek nesmí být velký

Každý předpis týkající se respondentů u jednotlivého, samostatného znaku i u kombinace znaků nazýváme kvótní podmínkou. Rozlišujeme kvótní podmínky tazatele a kvótní podmínky vcelku. V literatuře se uvádí počet 3—4 kvótních znaků pro tazatele; nejsou uváděny explicitní limity počtu hodnot a počtu kombinací znaků. Nepočítáme sem ty kvótní znaky, které neovlivňují práci tazatelů, například geograficko-správní umístění [je dáno umístěním tazatele].

Počet použitých kvótních znaků můžeme zvolit tak, že u jednotlivých tazatelů jsou respektovány určité konstantní podmínky, například tazatelé vybírají pouze buď ženy, nebo muže, a to navíc jen v jedné věkové skupině, přičemž dostávají kvóty pouze pro sociálně třídní určení a pro ekonomickou aktivitu; nebo naopak mají vybrat pouze muže-dělníky a kvóty dostávají vzhledem ke kvalifikaci a věku.[8] Tak je možnost zvýšit počet podmínek vcelku, neboť některé nezajišťuje tazatel. Je možné nejen přidávat kvótní znaky, ale přecházet ke kombinacím či volit podmínky speciální (počet svobodných matek, počet prodavaček ve velkých obchodních domech, počet dělníků pracujících v podzemí apod.)

3. Znaky i jejich hodnoty musí být zvoleny tak, aby byly snadno pochopitelné tazateli a respondentům:

- kategorie, které jsou jednoznačné a jednoznačně chápáné
- kategorie obecně používané a známé.

Jednou z chybných premis výzkumníků je, že stačí, aby byly kategorie kvótních znaků dobře definovány a v předpisech dobře popsány a že takové požadavky budou tazateli splněny, že se jim tazatelé obzvláště při opakovaném výzkumu naučí. To může být pravda u vyjimečných případů, obecně to však nemůžeme předpokládat. V průběhu času se rutinizuje spíše vlastní tazatelovo chápání názvu kategorie než (pro tazatele) nezvyklá definice výzkumníka. Kategorie budou dobře plnit svoji roli především tehdy, budou-li tazatelem pochopeny automaticky. To ovšem neznamená, že není nutné kategorie popisovat. Řadu kategorií je pochopitelně třeba upřesnit (dělník, zemědělec, zemědělský dělník, učeň, žena v domácnosti a na mateřské dovolené, vedoucí pracovník apod.). Musí však být voleny tak, aby odpovídaly všeobecnému povědomí a užívání. V předpisech tedy nebudeme používat vědecké pojmy, které se liší od normálního, běžně používaného jazyka, ale pojmy běžné řeči.

4. Kvalifikace respondentů podle kategorií kvótních znaků musí být pro tazatele jednoduchá, kategorie snadno identifikovatelné

Ověření vhodnosti respondenta pro kvótní předpis musí být nenásilné, jedno-

(7) Takový úkol však jistě není jednorázovou akcí. Nestačí tu hypotézy a apriorní subjektivní přesvědčení (tím lze začít), ani jednoduchý soubor empirických údajů. Jde o poznávací proces v celém širokém smyslu, sladění a postupné zařazování nových empirických poznatků do celkového fondu teoretických znalostí.

(8) Krajní případ, kdy tazatel nedostával žádné variabilní podmínky, ale pouze určené charakteristiky společné všem jeho respondentům, není metodologicky prozkoumán. Postup by však stál za prozkoumáním. Lze předpokládat, že má mnoho výhod i nevýhod; jejich vzájemný poměr však nelze pouhým metodologickým rozбором (bez empirického experimentu) stanovit.

duché a nesmí vzbuzovat neochotu odpovídat či nedůvěru ve výzkum nebo tazatele.

Typické důsledky porušení této zásady spočívají v tom, že některé kategorie (a pro výzkum mohou být velmi důležité) jsou podceněny (např. žena na mateřské dovolené, je-li kvótně řazena podle zaměstnání, které měla před nástupem na mateřskou dovolenou). Je také důvodem, proč mezi kvótními znaky nebývá například tak důležitá charakteristika, jako vzdělání. Zcela nevhodné je volit za kvótní znak příjmové skupiny: a) identifikační dotaz před zahájením by byl velice nevhodným vstupem, b) odpovědi nejsou solidní.

5. Dostupnost údajů pro stanovení kvót

U kvótních znaků musíme znát populační rozložení v kategoriích zvolených pro určení kvót, které nemá být zastaralé. Malé rozdílnosti způsobené vývojem však nemohou výsledky příliš zkreslit.

6. Nezávislost znaků

Kvótní znaky mají být statisticky nezávislé, aby celá soustava přinášela co nejvyšší určenost respondenta a co nejvyšší informaci o něm v co nejmenším počtu podmínek. Takové znaky, které jsou silně korelované, přinášejí vzájemně zastupitelnou informaci a tím se efektivita jejich zařazení částečně snižuje.

Uvedené požadavky na výběr kvótních znaků jsou dosti striktní a značně omezují naše možnosti a tím i možnosti kvótního výběru jakožto metody sběru reprezentativní informace. Určení kvótních znaků je složitá záležitost: působí zde faktory teoretického stavu disciplíny a souhrnu současných empirických poznatků, faktory praktických možností vzhledem k dostupnosti statistické informace, faktory vztahu obecně užívaných názvů a odborné terminologie (sociologické i statistické) i řada dalších praktických momentů.

B) Určení kvót

Již v první části stati byl kvótní výběr charakterizován jako postup, který má co nej přesněji zajistit shodu výběrového souboru a cílové populace z hlediska vybraných kvótních znaků, resp. jejich kombinací. Tato idea, která stála na počátcích vývoje metody, *může být rozšířena: volíme kvótní znaky, jejich hodnoty, zastoupení těchto hodnot (resp. kombinací hodnot) tak, aby dosažená přesnost výsledných odhadů byla co nejvyšší*. Volíme tedy zastoupení jednotlivých kategorií (skupin respondentů) z hlediska obecnějšího principu optimálního sběru dat, z hlediska širšího pohledu efektivity výzkumného procesu jako celku. Takovýto hlubší požadavek nemusí ovšem vést k proporcionálnímu určení kvót vzhledem k rozložení cílové populace, a to ze dvou možných důvodů:

- určení optimálních kvót vzhledem k co nejlepšímu podchycení variability populace
- zajištění dostatečně velkého podsouboru z nějaké málo zastoupené části populace, na kterou se koncentruje zájem výzkumníka (jde de facto o doplňkový výběr části populace).

První případ znamená „nerovnoměrné“, resp. neproportionální umístění výběrového souboru: nadhodnocuje kvóty u těch skupin obyvatelstva, u nichž je velká názorová variabilita, a podhodnocuje ty skupiny, u nichž je variabilita názorů, mínění a postojů nízká. Jde tu o podobný princip jako u stratifikovaného pravděpodobnostního výběru: Neymanova optimální alokace určuje výbě-

rovou velikost úměrně k součinu populačního zastoupení oblasti a směrodatné odchylky zkoumané proměnné.(9)

V praxi se objevují určité překážky pro jeho aplikaci:

— nepracujeme s jednou proměnnou, ale s mnoha, které si mohou vzájemně odporovat

— většinou směrodatné odchylky ani jejich poměry neznáme

Proto je aplikace Neymanova principu obtížná a předpokládá, jako i jiné nároky postupu, značné metodologicko-empiricko-teoretické úsilí.(10) Aplikace tohoto principu je však jednou z podstatných komponent umění kvótního výběru.

Pro určení kvót platí zásady, které mají úzkou souvislost se zásadami pro výběr kvótních znaků:

a) *Statistické podklady pro určení kvót by měly být relativně nové.*

Použitelnost a aktuálnost údajů závisí na znaku samotném, na jeho výzkumné roli a relevanci a na rychlosti změn v jeho populační distribuci. Některé údaje mohou být staré 3 roky (například prudký vzrůst žen na mateřské dovolené, rychlý vzrůst počtu rodin s více dětmi či naopak prudký pokles porodnosti, nástup populačně slabších ročníků ze základní školy apod.), některé údaje jsou použitelné po deseti či patnácti letech.

Podklady pro určení kvót jsou především:

- údaje celostátní statistické služby
- sčítání lidu
- mikrocensus — velká pravděpodobnostní šetření v celé populaci ČSSR
- statistické ročenky
- extrapolované údaje statistické služby
- velká účelová pravděpodobnostní šetření
- kontrolní pravděpodobnostní šetření
- zkušenosti z minulých kvótních i pravděpodobnostních šetření
- resortní statistiky, statistiky obvodů, odvětví, podniků apod.

Malé změny skutečné situace oproti dostupným údajům nemají na kvalitu kvótního výběru rozhodující význam.

b) *Rozmístění tazatelů* je současným určením kvót zastupujících:

- geografické rozmístění výběru
- administrativně správní rozmístění výběru
- velikost místa bydliště
- sociálně ekologické prostředí (velkoměsta, malá města, vesnice; zemědělské podniky, průmysl, školství, zdravotnictví, obchod atd.; hlediska sociální ekologie; typy ohrožení životního prostředí atd.)

c) *Určení místa rozhovoru* je také nástrojem určování kvót či s kvótami úzce souvisí. Rozhovor může být (podle instrukcí tazatelů) umístěn do různých prostředí, která chceme srovnávat, jako například

(9) Princip má jednoduchou interpretaci: čím větší je sociální skupina, tím více pozorování je třeba k jejímu pokrytí, ale současně: čím různorodější jsou názory ve skupině, tím více pozorování je třeba.

Princip lze ještě rozšířit o úvahy spojené s různými náklady v jednotlivých skupinách.

(10) Především jde o dlouhodobé registrace výsledků výzkumů a jejich hodnocení.

- různé části města
- různé zábavné akce, různé parky
- sportovní podniky
- společenská prostředí (restaurace, hospody, stánky)
- jednotlivé instituce určitého celku (různá ZOO).

Místo rozhovoru ovšem určujeme většinou jako homogenizační činitel pro prostředí, abychom nevnašeli do výzkumu nežádoucí obtížně kontrolovatelné faktory:

- bydliště respondentů
- pracoviště
- místa zábavy a oddechu
- veřejná místa (nádraží, ulice, park apod.)

U nás se většinou používá bydliště respondentů, v SSSR například jsou častěji respondenti vybíráni na pracovišti. Rozhodnutí ovšem závisí do velké míry na tématu rozhovoru.

d) *Kvótní podmínky jsou určovány specificky podle tazatelovy lokality, resp. podle místa rozhovoru* (obec, okres, město, závod, podnik apod.)

To znamená, že jsou stanoveny podle nejpodrobnějších statistik a údajů, které jsou k dispozici. Zvýšené úsilí při rozpisu kvót přináší vysoký metodologický zisk v kvalitě výběrového postupu a reprezentativitě šetření:

- zesiluje stratifikační efekt a potlačuje efekt vychýlení
- potlačuje umělou homogenizaci tazatelových respondentů a tím i skupinkový efekt
- ulehčuje práci tazatelům
- přispívá k přirozené volbě kvót vzhledem k prostředí

Formulovaný požadavek bývá z různých důvodů obcházen. Jako argument se většinou uvádí absence příslušných statistik. Přitom se však zapomíná na krajské a okresní statistiky a přehledy, které umožňují částečnou specifikaci vzhledem k prostředí, v němž tazatel provádí rozhovory. (11)

K určování proporcí patří ovšem i problematika určování kvót pro dílčí populaci, pro niž nejsou údaje k dispozici. V takovém případě bereme:

- nejlepší aproximaci, která je k dispozici
- údaje z jiných výzkumů
- odhady expertů
- metodu „sněhové koule“ (snowball) (12)

C) Výběr tazatelů a místo výběru

Výběr tazatelů pro jednotlivý výzkum se provádí ve dvou etapách

- 1) přijetí tazatelů ke spolupráci,
- 2) výběr konkrétních osob pro daný výzkum, tj. pro daný kvótní výběrový postup.

(11) Nedodržení zásady vede k tomu, že například pro zemědělský okres v rámci vysoce průmyslového kraje předepisujeme nevhodné kvóty (vysoké procento dělníků), nebo naopak respondenty z řad zemědělské populace vyžadujeme od tazatelů velkých měst.

(12) Při tomto postupu vyhledáváme další osoby podle údajů od již dotázaných respondentů. Metodu používáme především v případech, kdy jde o malé a dílčí populace rozptýlených jedinců, kteří mají však společné zájmy a alespoň částečně se mezi sebou znají.

Práce s tazateli je dlouhodobá záležitost a zahrnuje kromě jejich výběru i jejich výchovu, kontrolu práce, hodnocení, popřípadě zvlášť zaměřený permanentní výzkum. Celá problematika tazatelů by si zasloužila samostatnou stať shrnující zkušenosti uvedené v literatuře i zkušenosti našich pracovníků. Nebudeme se zde proto tématem zabývat nijak podrobně.⁽¹³⁾ O práci s tazateli, jak se provádí na ÚVVM při FSÚ, viz [Dvořák 1978].

Z hlediska výběru samotného existují dvě možné strategie, které připadají v úvahu:

1. Výběr tazatele je zároveň výběrem kvót z určitého prostředí a z určité geografické nebo správní oblasti. Tento přístup je většinou používán, podrobněji o něm byla už zmínka v zásadách pro určování kvót a kvótních znaků.

2. Výběr tazatelů může (či nemusí) být veden snahou o dodržení kvótních předpisů i u souboru tazatelů, aby tak byly eliminovány některé zdroje vychýlení.

D) Výběr respondentů tazatelem

O některých momentech výběru respondentů tazatelem jsme pojednali už při charakterizování kvótního výběru. Krok výběru respondenta tazatelem je *kritický pro úspěšnost celého postupu*. Záleží především na kvalitě a poctivosti tazatelské sítě, ale i na jasnosti instrukcí, na celkové práci s tazatelskou sítí. Již dříve jsme uvedli, že by bylo možné sestavit pravděpodobnostní model kvótního postupu za předpokladu, že by tazatelé vybírali v rámci předepsaných kvót respondenty náhodně (ve smyslu generování náhodných jevů). Takový výběr však nejen nelze zajistit, ale i pokud by se to podařilo, je vystaven rušivým vlivům.

Například na efekt vychýlení a na skupinkový efekt působí rušivé vlivy:

- a) preference určitého typu lidí tazatelem (preference může být buď záměrná, či neuvědomnělá);
- b) výběr osob z jednoho určitého města, z jednoho prostředí;
- c) tendence k vytváření vlastního panelu, u něhož tazatel pravidelně provádí dotazování, či k vytváření okruhu respondentů, z nichž si cyklicky vybírá;
- d) nepoctivost či chybnost při určování a záznamu kvótních znaků;
- e) nepoctivost při vyplňování dotazníku (tazatel imituje některé respondenty — vymýšlí si sám odpovědi);
- f) tazatel vybírá (bez ohledu na odmítnutí) tak dlouho, až splní kvóty, takže se tu projevuje efekt vychýlení jako výsledek mechanismů samovýběru a odmítnutí;

(13) Přestože tazatel je základním kamenem úspěšnosti celého procesu sběru dat, víme o tomto problému zatím velmi málo. K řešení podobných úloh by mohlo především být využito možnosti pravidelného opakování výzkumů založených na kvótním výběru. Pracoviště, které pravidelně organizují výzkumy, mohou při nepatrných nákladech a bez jakéhokoli rizika takovato vyhodnocení provádět ve spojení s drobným metodologickým experimentováním (které však ani nekomplikuje organizaci, ani nezvyšuje náklady). Zde je podle našeho názoru pole pro velmi podstatnou metodologickou práci a tím pro zlepšení kvality kvótního výběru v našich podmínkách, a tedy pro zvýšení efektivity a kvality našich sociologických šetření, která jsou na této metodice závislá.

g) tazatel vybírá osoby, které jsou přímo dosažitelné, vynechává osoby, které nejsou doma, mají málo času v okamžiku kontaktování, jsou blízko tazatelova bydliště apod., a tím redukuje cílovou populaci na její snadno dosažitelnou část;

h) tazatel není schopen po snadném výběru prvních respondentů dodržet kvótní podmínky u posledních dotazovaných — to vede k porušení kvót (připravenému či nepřipravenému) či ke snížení počtu respondentů.

Vliv tazatele na reprezentativnost výběrových dat je nesporný. Jeho osobnost, postoj k výzkumu, poctivost a svědomitost mají kritický význam. Je-li tazatel schopen vhodně navázat kontakt, vykazuje menší počet odmítnutí a dostává solidnější odpovědi. Práce s tazatelem je náročná činnost, patří k ní perfektně vypracované manuály (hezky vypravené a ilustrované), pravidelná setkání, školení, pravidelné informace, kontroly, vhodné metody hodnocení.(14) K základním metodám práce s tazateli patří také postupné zvyšování náročnosti a pravidelná obměna tazatelských kádřů.(15) Je ovšem pochopitelné, že jinak pracujeme u dlouhodobých tazatelských sítí a jinak v případě sestavení ad hoc tazatelské sítě k jednomu výzkumu.

E) Zhodnocení reprezentativity

je obtížný problém, chápeme-li reprezentativitu komplexně a v plné její šíři (viz Řehák 1978). Jednotlivé metody hodnocení reprezentativity vešce záleží na konkrétních výzkumných podmínkách. Její ověřování u kvótního výběru je komplikováno absencí matematického modelu. Musíme proto nutně zkoumat chování metody empiricky a pokusit se sestavit model indukční metodou založenou na kontinuální zkušenosti nebo na základě předem připravených ověřovacích postupů.

O statistickém ověřování reprezentativity pojednává Řeháková (1979), (1980). Hodnotit reprezentativitu je nutné, neboť jen tak můžeme identifikovat zdroje vychýlení a nalézt prostředky k jejich redukci.

F) Analýza dat

Znalost modelu chování dat a vyhodnocení reprezentativity jsou nutnými podmínkami pro určení způsobu analýzy dat(16) Metody analýzy dat jsou při-

(14) O nevhodných formálních ukazatelích práce tazatele (i tazatelské sítě) byla už řeč. Zde je však nutno poznamenat že hodnocení tazatele nesmí poškozovat ty, kteří se otevřeně a poctivě přiznají k chybám, nedostatkům, nedodržení kvótám, a to v relaci k těm, kteří mají všechny ukazatele v pořádku bez ohledu na způsob dosažení tak vysoké „kvality“. Odlišení kvalitní práce od nepoctivého zjednodušování situace je nesmírně obtížné. Ale je přesto nutné vyvíjet vhodné ukazatele. Automatizované zpracování a ukládání dat k tomu otevírá nové možnosti: statistické zpracování údajů sebraných jednotlivými tazateli v různých výzkumech, kontrola variability mezi tazateli ve stejných prostředích a se stejnými kvótními předpisy apod. K tomu je ovšem nutno vytvořit předpoklady výzkumem práce tazatelů, hledáním modelů procesu, kterým tazatelé vybírají respondenty atd.

(15) V žádném případě nelze přijmout postoj: „buďme rádi, že máme potřebný počet tazatelů; nemůžeme jim dávat příliš obtížné úkoly, aby s námi nepřestali spolupracovat.“ S takovýmto (bohužel u nás často v nějaké formě existujícím) přístupem nelze vybudovat dobrou tazatelskou síť a nelze pak ani dělat dobré a reprezentativní výzkumy se solidními a spolehlivými závěry.

(16) Opomeneme-li tyto dvě nutné podmínky, pak nemáme vůbec žádné opodstatnění používat běžné metody statistiky: chí-kvadrát test, konfidenční intervaly, bodový odhad, tedy vše co běžně používáme ve výzkumných zprávách. Argumenty založené na podobných postupech jsou pak neopodstatněné, a tudíž vědecky neplatné.

mo svázány se způsobem sběru a jsou závislé na různých faktorech, které přitom působí. Úkolem analýzy dat je: charakterizovat nepřesnosti a náhodné chyby výběru i měření a eliminovat vychýlení při odhadech populačních parametrů charakterizujících stav u jednotlivých proměnných i vztahy proměnných navzájem, a to pro celou populaci i její vybrané části; resp. vybrané části populace komparovat. Úkolem analýzy dat je tedy provést zobecnění získaných chybových dat na danou cílovou populaci a daný výsek sociální reality, který je předmětem výzkumu. Abychom mohli (bez formálního platného modelu) charakterizovat náhodnou chybu výběru, provádíme (při dlouhodobém opakování kvótního výběru) porovnání s předpokladem prostého náhodného výběru. Kvótní výběry obvykle zpracováváme jako prosté náhodné výběry. K tomu vede úvaha, že proti sobě stojí dva opačně orientované efekty: stratifikační a skupinkový, které se ruší. Údaje o platnosti takového předpokladu se v literatuře dosti liší. Moses a Stuart (1953) považují předpoklad za nepřijatelný, Stephan a Mc Carthy (1958) uvádějí analýzu vybraných dat, která též spíše vede k odmítnutí hypotézy o empirické ekvivalenci chyb kvótního a prostého náhodného výběru; podrobnější analýza těchto údajů však plně odpovídá takové hypotéze. Konečné závěry je nutno vztáhnout ke konkrétnímu pracovišti, jeho tazatelské síti, organizaci, podmínkám.

Mluvíme-li o vztahu výběrových chyb dvou typů výběrového uspořádání, máme na mysli porovnání vzhledem k vychýlené populaci. Náhodné chyby nám nepodávají informaci o systematickém vychýlení (viz Řehák 1978). (17)

Jasně implikace pro zpracování a analýzu dat máme jen tam, kde jsme sami vnesli do výběru formalizovaný záměr: tam kde jsme volili neproporcionální kvóty (například pro dosažení maximální výběrové efektivity).

Můžeme se snažit podchytit a definovat i řadu dalších faktorů vychýlení, to už však opět patří k onomu umění kvótního výběru. Vychýlení je velmi nepříjemný a obtížně identifikovatelný efekt nedůslednosti a nekontrolovatelnosti metody. Je však nutné se s ním vypořádat — na tom závisí životaschopnost celé metody.

Zkoumání náhodného chování odhadů kvótního výběru vzhledem k prostému náhodnému výběru je důležité i tehdy, jestliže předpoklad o empirické ekvivalenci chování obou je porušen. V tom případě je třeba znát poměry směrodatných chyb a použít tuto informaci při statistické inferenci z dat.

G) Metodologické vyhodnocení zkušeností z výzkumu a závěry

Tento krok dává kvótnímu výběru charakter neustálého procesu a reflexe všech zkušeností pro další výzkumné šetření. Největším uměním v kvótním výběru je umění využít všech předchozích informací a zkušeností z již provedených šetření pro další následné výzkumy, tj. umění neustále modifikovat postup, varovat se zjištěných chyb, plně využít dobrých zkušeností, a tak se blížit k aproximaci optima.

(17) Občas se vyskytuje otázka, co je důležitější, zda vychýlení či nemožnost odhadu, či zda má vůbec smysl počítat neurčitost odhadu kolem vychýleného odhadu. To souvisí se zcela konkrétními úlohami. Pro popis populace nutno nejprve redukovat vychýlení a pak určit neurčitost odhadu. Pro testy dobré shody je nutné znát neurčitost. Pro komparace jednotlivých částí populace nemusí vychýlení vadit za předpokladu, že působí na všech částech stejným směrem a se stejnou intenzitou. Obě charakteristiky přesnosti jsou důležité.

3. Závěr

Cílem této stati není ani prosazovat postup kvótního výběru, ani shrnout argumenty proti jeho používání. Kvótní výběr je používán především proto, že je považován za organizačně jednodušší i za méně nákladný.⁽¹⁸⁾ Jako každá metoda, i kvótní výběr má své přednosti a nedostatky. I když není možné shrnout celý postup do jednotné metodiky, přesto vidíme, že jde o metodu stejně náročnou jako jiné metody. Neexistuje bohužel opravdu úplné a systematické srovnání kvótního výběru a ostatních metod výběru, takže nevíme, zda úsilí, které je nutné pro překonání kritických bodů, není vyšší než úsilí nutné pro přípravu pravděpodobnostního výběru.

Kvótní výběr je metoda natolik spjatá se specifickým použitím, organizací i sociálním prostředím či situací, v níž je aplikována, i s tématem výzkumu, že otázky naznačené v této stati musí být zodpovězeny (jako všechny ostatní otázky související s kvótním výběrem) dlouhodobou praxí a neustálým vyhlazováním postupu. Zkušenosti s tímto postupem jsou publikovány v odborné literatuře celkem sporadicky a nelze je převzít, protože neodpovídají konkrétní situaci našich výzkumů. Mohou sloužit jako podklady metodologických hypotéz a jako orientace ve složité situaci. Nutně však musíme spoléhat na svůj vlastní metodologický výzkum, na metodologická pracoviště a zobecnění vlastních zkušeností. Cílem stati bylo podat rozbor postupu a ukázat na jeho různé aspekty, s nimiž je nutno se v práci vypořádat. Nejde o metodu jednoduchou a bezespornou (jak se řada výzkumných pracovníků domnívá). Je-li používána, aniž by byla průběžně metodologicky zdokonalována, je výrazem nekvalifikovaného a nezodpovědného přístupu k vědecké a výzkumné práci i k zadavateli.

Literatura

- Dvořák, P.: *K práci s tazateli ve výzkumu veřejného mínění*. Sociologický časopis 1978, č. 6, s. 644–656.
- Katz, D.: *An Analysis of the 1948 Polling Predictions*. Journal of Applied Psychology 1949, No. 33, pp. 15–28.
- Kish, L.: *Survey Sampling*. New York, Wiley 1965.
- Moses, C. A. — Stuart, A.: *An Experimental Study of Quota Sampling*. Journal of the Royal Statistical Society (A) 116, 1953, pp. 349–405.
- Řehák, J.: *K pojmu „reprezentativita“ v sociologických výzkumech*. Sociologický časopis 1978, č. 6, s. 489–507.
- Řeháková, B.: *Statistické ověřování reprezentativity: testy dobré shody*. Sociologický časopis 1979.
- Řeháková, B.: *Statistické ověřování reprezentativity: koeficienty neshody*. Sociologický časopis 1980 — v tisku.
- Stephan, F. F. — McCarthy P. J.: *Sampling Opinions*. New York, Wiley 1958.

(18) Taková charakteristika je většinou uváděna v literatuře i na odborných diskusích. Tento názor je však intuitivní a v našich podmínkách neprokázaný. Nákladnost metody je nadto relativní vzhledem k přesnosti výsledků a k validitě interpretačních závěrů.

Резюме

Я. Ржегак: Квотная выборка

В статье описывается метод квотной выборки в её связи с исследовательскими задачами. Квотная выборка является комплексом различных техник и методов, которых надо приспособить к специфическим и конкретным условиям данной организации, данного исследования и данной социальной ситуации.

В первой части приводится характеристика квотной выборки и описаны три типа эффектов, которые определяют качество методологического применения и которые в дальнейшем являются основными критериями методологической дискуссии метода вообще и его отдельных этапов: эффект районирования, эффект группировки и эффект смещения. В этой части даётся список самых важных факторов влияющих на эти три вида эффектов и механизмы их деятельности.

Вторая часть статьи приносит список семи этапов работы при квотной выборке:

- А) определение квотных переменных
- Б) определение квотных правил для выборки
- В) подбор спрашивающих
- Г) подбор респондентов
- Д) оценка репрезентативности
- Е) анализ данных
- Ж) методологическая оценка накопленного опыта

Отдельные этапы разбираются методологически и у них даны методологические правила для применения на практике.

Целью статьи нет снести аргументы для или против применений квотной выборки. Автору хотелось описывать метод со всеми его преимуществами и препятствиями и так дать в руки исследователей информацию и основы для самостоятельного решения о применимости метода в его собственных условиях.

Summary

J. Řehák: Quota Sample

The paper describes the method of quota sample in its broad sense. The quota sample is regarded as a whole complex of various methods and procedures which have to be specified to particular conditions of an organization, research and social situation.

The first part presents a general characteristics of quota sample. Three types of effects are introduced: effect of stratification, clustering effect and effect of bias. These are then used in the following methodological analysis. The main factors influencing these three factors are described.

In the second part, the seven steps of the sampling process are described and methodologically discussed:

- A. choice of quota variables
- B. prescription of quotas
- C. selection of interviewers
- D. selection of respondents
- E. checking of representativity
- F. data analysis
- G. methodological evaluation of the overall experience

Methodological analysis of all these steps as well as rules for practical use are given.

The aim of the paper is not to give an argument for or against the method of quota sample. It presents all its advantages and dangers and leaves it to the user to decide whether it can be used in a particular situation. The applicability and modifications are always subject to special conditions.