

AZ ÁLTALÁNOS CÉLÚ HÁZTARTÁSI MINTA KIALAKÍTÁSA

A Magyar Tudományos Akadémia Statisztikai Bizottsága és Demográfiai Bizottsága 1972. április 27-én együttes vitaülést tartott. Az ülésen a két bizottság részéről Huszár István államtitkár, a Központi Statisztikai Hivatal elnöke és dr. Szabady Egon, a Központi Statisztikai Hivatal elnökhelyettese elnökölt. A meghívottaknak előre rendelkezésre bocsátották az alább közölt vitaindító összefoglalás anyagát, melyet az ülés előadói, dr. Bene Lajos, a Központi Statisztikai Hivatal Népszámlálási osztályának munkatársa és Éltes Ödön, a Központi Statisztikai Hivatal osztályvezető-helyettese állítottak össze.

I. A MINTAVÉTELI MÓDSZEREK ÖSSZEFOGLALÓ ISMERTETÉSE

A Központi Statisztikai Hivatalnak a közeljövőben két nagyobb mintavételes feladatot kell megoldania. Az egyik az 1973. január 1-i időpontra tervezett mikro-cenzus és a hozzá kapcsolódó jövedelmi felvétel módszertani és technikai megvalósítása. A másik feladat olyan mintavételi keret létesítésében áll, amely a következő években lebonyolítandó különböző képviseleti adatgyűjtéseket szolgálná. Ilyen adatgyűjtések:

- a munkaerő összetételének folyamatos megfigyelése (munkaerőmérleg),
- háztartásstatisztikai adatgyűjtés,
- különböző társadalomstatisztikai, egészségügyi, kulturális stb. felvételek.

E célra szükséges és kívánatos a Központi Statisztikai Hivatalban olyan állandóan működő egységes háztartási adatgyűjtési rendszer létrehozása, amely állandó szervezettel, személyi megkérdezésre kiképzett összeírókkal és egy állandóan frissen tartott mintavételi kerettel rendelkezik. Ez a szervezet az eddiginél magasabb szinten tudna eleget tenni a jelenlegi differenciált adatigényeknek és megfelelne a jövőben várható növekvő követelményeknek is.

Az egységes háztartási adatfelvételi rendszer főbb előnyei a következők:

1. kiküszöbölhető a tematikai átfedés;
2. csökkenthető a lakosság igénybevétele a felvételek témáinak koordinálása által;
3. egységes szervezet gondoskodik a felvételek szervezéséről és végrehajtásáról, állandó begyakorolt és rendszeresen oktatott összeíró gárda útján;
4. egyik népszámlálástól a másikig folyamatosan, naprakész állapotban tartott minta áll az egyes adatfelvételek rendelkezésére, mely matematikai statisztikai megbízhatóság szempontjából is a legelőnyösebben alakítható ki.

Az egységes háztartási adatfelvételi rendszer kialakítására az érintett közgazdasági, népesedésszatisztikai és társadalomstatisztikai főosztályok közreműködésével már megkezdődött a munka, mégpedig munkabizottságok útján.

Az egységes háztartási felvételi rendszer sokrétű feladatait ún. általános célú mintával kívánjuk megoldani. Ez a minta területi egységeket tartalmazó „elsődleges” mintavételi egységekre vonatkozik úgy, hogy a kiválasztott elsődleges egységekbe tartozó megfigyelési egységek – háztartások, személyek, esetleg más egységek – hivatva vannak az egész országot reprezentálni. Az általános célú minta elsődleges egységei által felölelt sokaságból történik az egyes konkrét felméréseknél alkalmazott megfigyelési egységekből álló minta kiválasztása.

Az általános célú minta számos előnnyel rendelkezik.

Mindenekelőtt egyszerűbbé teszi a mintavételt, mert az elsődleges mintavételi egységek alapos előkészítő munkát igénylő kiválasztását csak egyszer kell végrehajtani.

Ezenkívül a megfigyelési egységeket tartalmazó listát, a mintavételi keretet csak a kiválasztott és hosszabb időszakon keresztül változatlan elsődleges mintavételi egységeken belül kell kompletten tartani.

Végül a hibaszámítást is egyszerűbbé teszi az általános célú minta alkalmazása, ugyanis a kiválasztásból eredő hiba egyik – sok esetben döntő – része, az elsődleges egységek közötti szórásnégyzet, számos ismételten vonatkozólag előre meghatározható, s csak az elsődleges egységeken belüli kiválasztásból eredő hibátags becslését kell az egyes konkrét mintavételeknél meghatározni.

Az általános célú minta két népszámlálás közötti időszakban szolgál keretül az egyes speciális mintavételekhez. A következő népszámlálás feldolgozása után, illetve azzal párhuzamosan új általános célú minta kerül kiválasztásra, amely újabb tíz évre szolgál mintavételi keretül.

Ezen általános célú minta legcélszerűbb mintavételi egysége a népszámlálási számlálókörzet. Az összeírás végrehajtása és a címjegyzékeknek hosszabb időn át naprakész állapotban való tartása szempontjából ugyanis a területi minta előnyei kétségtelenek.

Előzetes megfontolások szerint az általános célú mintának a népszámlálási körzetek kb. 10 százalékát kell felölelnie, pontosabban annyi számlálókörzetet, hogy a mintába került körzetek népessége az ország lakosságának 10 százalékát képviselje. Ez a mintanagyság kb. 335 000 háztartást, ezekben mintegy 1 030 000 személyt jelent.

A tervezett általános célú minta felépítése kétlépcsős rétegzett kiválasztással oldható meg: az első lépcsőben községek, a másodikban ezeken belül számlálókörzetek kiválasztásával.

A községek kiválasztása a minta területi tömörítése érdekében, a helyszíni munka gyakorlati követelményei szempontjából szükséges.

A lakosság szám nagyság bizonyos határán felüli települések azonban mind mintába kerülnek: itt szisztematikus rétegzett számlálókörzeti kiválasztás történik. Kiindulópontul az az elgondolás szolgál, hogy az említett lakosság szám nagyság-határ, amelyen felül minden település mintába kerül (az 1963. évi és az 1968. évi mikrocenzusok bevált módszerével egyezően) 5000 lakos. Az e határnál kisebb községekből általában községenként 2–2 számlálókörzetet veszünk mintába.

Az általános célú minta teljes számlálókörzeteket tartalmaz, míg a belőle képzendő kisebb minták már e terület egységeken belül háztartások vagy lakások kiválasztásával készülnek. A mikrocenzus tervezett 2 százalékos mintája még valamennyi terület egységre kiterjed, a folyamatos adatgyűjtéseket szolgáló kisebb minták azonban majd részben a községek, illetve számlálókörzetek számának további csökkentésével, részben ezekben a háztartások számának megfelelő arányú kiválasztásával készülnek.

A minta tényleges elkészítéséhez az 1970. évi népszámlálás adja meg a szükséges alapot. Mégpedig

1. szolgáltatja a listákhoz szükséges címanyagot,
2. még ezt megelőzően a minta felépítéséhez, rétegzéséhez nélkülözhetetlen területi részletezésű adatokat biztosítja.

A népszámlálási adatok alapján történő gondos felépítésre, vagyis a rétegzésre nagy súlyt kell helyezni, mert ezzel egyenlíthetők ki a területi mintának az egyedi vagy a háztartási mintával szemben jelentkező nagyobb hibahatárai.

Az általános célú minta több éven keresztül fog mintavételi keretül szolgálni számos felvételhez, ezért a lehető legmesszebbmenőkig biztosítani kell, hogy minden szempontból valóban reprezentálja az ország népességét.

Az optimális rétegzés megoldása nem egyszerű feladat. Először is a többlépcsős kiválasztásnak megfelelően a rétegzés szükségessége nemcsak a községek, hanem a számlálókörzetek vonatkozásában is felvetődik. Másodsorú célszerű ugyan a minél mélyebb rétegzés, azonban minél mélyebben rétegzünk, annál nehezebb biztosítani, hogy a minta elosztása ugyanakkor arányos is legyen. Ezért a mintavétel gyakorlati végrehajtása nem képzelhető el bizonyos kompromisszumok, illetőleg utólagos korrekciók nélkül.

1. A mikrocenzus mintájának felépítése

Az 1963. és az 1968. évi mikrocenzusok mintája – az akkor kitűzött pontossági követelményeknek megfelelően – 2 százalékos volt. Ez a kiválasztási arány mind a megbízhatósági igények, mind pedig a szervezet és költségvetés szempontjából megfelelőnek bizonyult és így azt a következő mikrocenzusnál is alkalmazni kívánjuk.

Területi tömörítése ugyanaz, mint az általános célú mintánál, tehát az ott kiválasztott valamennyi község és számlálókörzet a 2 százalékos mintában is szerepel. Az általános minta valamennyi községének és számlálókörzetének mintában tartása két szempontból indokolt:

1. a megbízhatósági határok szempontjából: a csoportos (területi) kiválasztás úgyis feltehetőleg csökkenti a minta hatékonyságát, ezért lehetőleg minél több területi egység bekapcsolása indokolt,
2. az általános célú minta címjegyzékének ellenőrzése és kiegészítése érdekében a mikrocenzusok alkalmával előirányozzuk az általános célú minta teljes számlálókörzeteinek bejárását.

A 2 százalékos kiválasztási arány azt jelenti, hogy a mikrocenzus alkalmával az általános célú minta $\frac{1}{5}$ részére kell a terjedelmet redukálni. A minta jósága érdekében az általános célú minta valamennyi területegységét mintában tartva, a kiválasztott számlálókörzeteken belül minden ötödik lakás (vagy háztartás) kerül be a mikrocenzus mintájába, azaz átlagosan 80–100 háztartást számolva körzetenként, 16–20 háztartás egy-egy körzetből. A mikrocenzus mintája kb. 67 000 háztartást, 206 000 személyt foglal magában.

A tervezések során felmerült a számlálókörzeteken belül zárt területi egységek kialakítása (mint az 1970. évi népszámlálás utóellenőrzésénél $\frac{1}{4}$ vagy $\frac{1}{5}$ számlálókörzet) és megvizsgáltuk ennek előnyeit a számlálókörzeten végigmenő lakás-kiválasztással szemben. Úgy láttuk azonban, hogy a számlálókörzetek esetleges inhomogenitására való tekintettel a számlálókörzet egész területét felölelő lakás-kiválasztás jobb mintát eredményez.

A mikrocenzus mintája tehát 5000 lakosnál kisebb községekben három-, ennél nagyobb községekben és városokban kétlépcsős.

A kiválasztás módjára, tehát a minta felépítésére többféle eljárás vehető igénybe, aszerint, hogy a községek nagyságukkal (a lakosok vagy a lakások számával) arányos, vagy egyenlő valószínűséggel kerülnek-e kiválasztásra. Ugyanígy a számlálókörzetek kiválasztása is elvégezhető nagyságukkal arányos vagy egyenlő valószínűséggel. A községek, illetve számlálókörzetek kiválasztási módszerének megválasztása azonban – a II. fejezetben előírt alapelv, illetve követelmény folytán – már eldönti a harmadik lépcsőben a lakások kiválasztásának módját.

Az összefoglaló ismertetéshez kapcsolódva bemutatjuk a különböző mintavételi lehetőségek megfelelő voltát, az ott feltüntetett követelmények, illetve elvek alapulvételeivel. A tényleges megvalósításhoz a II. fejezetben felsoroltak közül az 1. megoldást választottuk. Eszerint a községek és a kiválasztott községeken belül a számlálókörzetek mintába vétele a lakásszámukkal arányos valószínűséggel történik: minden körzetből a harmadik lépcsőben egységesen 17 lakás, egy-egy mintába vett 5000 lakosnál kisebb községből pedig 34 lakás kerül mintába. (Ezek a számok az 1970. évi népszámlálás címjegyzékére vonatkoznak, az új építkezések ezen felül lépnek a minta keretébe.)

Községenként azonos számú lakásból álló minta a számlálóbiztosok egységes munkaterhelése szempontjából előnyös.

A kiválasztás részletesebb ismertetését az alábbiak tartalmazzák.

Meg kell még említeni, hogy a mintavételi terv megalapozásához nélkülözhetetlen a két-, illetve háromlépcsős (község – számlálókörzet – lakás) minta megbízhatósági határainak számítása. A tervezett mintanagyságok korábbi tapasztalatokon nyugvó megfontolásokból adódtak, míg a többlépcsős minta hibahatárára vonatkozólag rendelkezésre álló támpontok eléggé szűk körűek. A mintaképzés kapcsán most a gépi feldolgozás szolgáltatja a számlálókörzeti átlagok szórására vonatkozó adatokat kb. 20 legfontosabb demográfiai és foglalkozási ismérvre és ezekből az ellenőrző számítások folyamatban vannak. Az 1970. évi népszámlálás utóellenőrzésének feldolgozása egyébként adatokat fog szolgáltatni 400 teljes és ugyanennyi egynegyed számlálókörzet anyagából nagyobb számú ismérv mintavételi hibája tekintetében is. Számításokat tervezünk ezenkívül utólagosan, a mikrocenzus és a jövedelmi felvétel mintájának több független részmintára bontása útján.

2. A községek kiválasztása

A községek nagyságával (esetünkben a községben levő lakások számával) arányos valószínűségű kiválasztást – hasonlóan, mint az előző mikrocenzusnál – kumulált lakásszámmal, véletlen kezdőszámmal és 1700 lakásos váltószámmal valósítjuk meg, kiválasztva azokat a községeket, amelyekre a kumulált sorban ez a váltószám esik. (A 2 százalékos mintának 50-es szorzószám felel meg, községenként 34 lakás, tehát $50 \times 34 = 1700$.)

A községeket népességnagyság-csoportok szerint rétegezzük, a népszámlálási kiadványokban is alkalmazott alábbi rétegek szerint:

– 499 lakos
500– 999 lakos
1000–1499 lakos
1500–1999 lakos
2000–2999 lakos
3000–4999 lakos
5000– lakos

A községek kiválasztására csak az 5000 lakosnál kisebb községeknél kerül sor, mert az ennél nagyobbak mind szerepelnek a mintában. A minta összesen kb. 1000 községet fog tartalmazni.

Ezen belül foglalkozási rétegzést alkalmazunk, mégpedig az 1963. és az 1968. évi mikrocenzusoknál használt rétegekkel: községek 40 százaléknál kisebb, 40–59,9 százalékos és 60 százaléknál több mezőgazdasági népességgel. A községminta elkészülte után ellenőrizzük a kiválasztás helyességét e rétegek, továbbá a járások arányos képvisellete szerint, és számottevő eltérés esetén azt községek cseréjével korrigáljuk.

A rétegzés kapcsán vita tárgya volt, hogy szükséges-e a községek foglalkozási rétegzése, tekintve, hogy a második lépcsőben a számlálókörzeteket úgyis rétegezzük. A döntés végül a fenti megoldás értelmében történt.

Olyan községekhez, amelyekben nincs két használható számlálókörzet (500 lakoson aluliak között) még egy hasonló, földrajzilag közelfekvő kis községet választunk ki.

A minta korrekciójához tartalék községeket is kiválasztunk, a minta 5 százalékának megfelelő mértékben.

3. A számlálókörzetek kiválasztása

A számlálókörzetek kiválasztása – éppúgy mint a községeké – nagyságukkal arányos valószínűség mellett, községenként történik. A minta felépítési terve szerint 5000 lakoson aluli községekben egységesen községenként két számlálókörzetet veszünk mintába, 5000 lakoson felül pedig a számlálókörzeteknek kb. 10 százalékát választjuk ki.

A kiválasztási művelet 5000 lakosnál kisebb községeken belül a számlálókörzetek számozásának rendjében halad, véletlen kezdőszámmal és az összes lakás-szám felének megfelelő váltószámmal. Az így kiválasztott számlálókörzeteket azután megyénként ellenőrizzük azok foglalkozási rétegzése szerint, összehasonlítva az összes és a kiválasztott számlálókörzetek rétegek szerinti megoszlását. A szükséges korrekciót községen belül a számlálókörzetek cseréjével hajtjuk végre.

Az itt alkalmazott foglalkozási rétegek a következők. Belterületen:

- 70 százalékon felüli ipari, építőipari aktív kereső népességgel,
- 50–70 százalékos ipari, építőipari aktív kereső népességgel,
- 50 százalékon aluli, de relatív többségben levő ipari, építőipari aktív kereső népességgel,
- 70 százalékon felüli mezőgazdasági aktív kereső népességgel,
- 50–70 százalékos mezőgazdasági aktív kereső népességgel,
- 50 százalékon aluli, de relatív többségben levő mezőgazdasági aktív kereső népességgel,
- 70 százalékon felüli egyéb aktív kereső népességgel,
- 50–70 százalékos egyéb aktív kereső népességgel,
- 50 százalékon aluli, de relatív többségben levő egyéb aktív kereső népességgel.

A külterületeken csak 2 réteg szerepel:

- mezőgazdasági többségű aktív kereső népességgel,
- nem mezőgazdasági többségű aktív kereső népességgel.

Külön kezeljük a 100 főn aluli kis számlálókörzeteket, mert ezekből a kellő számú lakás nem választható ki. Ezeket a 10 százalékos mintának megfelelő arányos számban azonos típusú, szomszédos, mintába vett számlálókörzettel vonjuk egybe. A kizárólag intézeti háztartásokból álló számlálókörzetek nem vesznek részt a kiválasztásban, ezekre külön eljárást alkalmazunk.

Az 5000 lakosnál népesebb községekben és városokban a számlálókörzeteket foglalkozási rétegek szerint csoportosítjuk és így végezzük a lakásszámok kumulációját, majd 850-es váltószámmal (számlálókörzetenként 17 lakás \times 50) a kiválasztást.

Felmerült olyan szempont, hogy városokban a foglalkozás mellett nagy fontosságú a földrajzi (területi) rétegzés, vagyis a város belső és külső részeinek arányos képviselete. Ezt pedig leginkább a számlálókörzetek sorszáma szerinti kiválasztás biztosítaná. Kísérleti munkák tanulsága szerint a földrajzi arányos képviselet foglalkozás rétegzés mellett is megfelelően érvényre juttatható.

4. A lakások kijelölése

Az 1970. évi lakásállomány (3 150 000) 2 százaléka 63 000 lakás. Így a kb. 40 000 számlálókörzet, illetőleg az intézeti és a 100 főnél kisebb körzetek levonásával kb. 37 000 körzet 10 százalékából, 3700-ból ($63\,000 : 3700 =$) 17 lakást kell számlálókörzetenként mintába venni.

A mikrocenzussal összeírandó lakások kijelölését az 1970. évi népszámlálás számlálókörzeti gyűjtőívén végzik.

A kijelölés véletlen kezdőszámmal és a 17 lakás alapján számítandó váltószámmal, mechanikus lépésekkel történik.

Egyébként az egész számlálókörzet előzetes ellenőrzésére tekintettel a körzet teljes címjegyzéke lemásolásra kerül. Ezen kell kijelölni a mikrocenzus kapcsán összeírandó 17 lakást, továbbá ezek közül a jövedelmi és az átrétegződési felvételben résztvevő lakásokat.

5. A jövedelmi és az átrétegződési felvétel mintája

A jövedelmi felvétel egyik feladata, hogy nemcsak országosan, hanem megyénként is megbízható adatokat szolgáltatson legalább a főbb népességcsoportok jövedelmi helyzetéről. Az 1968. évi jövedelmi felvétel megyei adataira vonatkozó hibaszámítások tanúsága szerint ahhoz, hogy a megyékre a kívánt mélységben kapassunk megbízható adatokat, kb. akkora mintára van szükség egy-egy megyéből, mint amekkora 0,5 százalékos kiválasztás mellett a legnagyobb megyékre jut.

Ez a követelmény megyénként különböző kiválasztási aránnyal biztosítható. Az előzetes számítások szerint négy különböző kiválasztási arány alkalmazására lett volna szükség: 0,5, 0,75, 1 és 1,5 százalékosra. A kiválasztás és a feldolgozás egyszerűbbé tétele érdekében a jövedelmi felvételnél négy megye és Budapest kivételével egységesen 1 százalékos mintát kívánunk alkalmazni. Bács-Kiskun, Borsod-Abaúj-Zemplén, Hajdú-Bihar és Pest megyék, valamint Budapest esetében a jövedelmi minta a mikrocenzus anyagának csak 25 százalékát fogja felölelni, azaz csak a háztartások 0,5 százalékára terjed ki. A megyei városok megyéjük mintájához igazodnak (Miskolc és Debrecen tehát 0,5 százalékos).

Előzetes számítások szerint ily módon a jövedelmi minta 26–27 000 háztartást fog magában foglalni.

Azokban a megyékben, ahol a jövedelmi minta 1 százalékos, a kiválasztás a mikrocenzusi mintából a lakások szintjén történik, azaz minden mintába került körzetben minden második lakás alkotja a jövedelmi mintát. Budapesten, Miskolcon, valamint Debrecenben a jövedelmi minta kiválasztása a mikrocenzusi mintából két lépcsőben történik, minden második körzet minden második lakása kerül bele a mintába. Ugyancsak kétlépcsős a jövedelmi minta kiválasztása Borsod-Abaúj-

Zemplén és Pest megyék 10 000-nél kisebb lélekszámú községeire, itt a mikrocenzusi mintában szereplő községek felében minden második lakásban fog sor kerülni a jövedelmi felvétel végrehajtására. Bács-Kiskun és Hajdú-Bihar megyében, valamint Borsod-Abaúj-Zemplén és Pest megyék 10 000-nél nagyobb lélekszámú településeiben (Miskolc és Debrecen kivételével) a kiválasztás egylépcsős: minden körzetből minden negyedik lakás alkotja a jövedelmi mintát.

A jövedelmi felvétel adataira – előző gyakorlatunknak megfelelően – a független véletlen részminták módszere alapján kívánunk hibaszámítást végezni. Ennek érdekében a jövedelmi minta öt független, azonos eljárással kiválasztott részmintából tevődik össze. Hasonló részminták képezhetők a mikrocenzus anyagából is.

A mikrocenzushoz foglalkozási átrétegződési felvétel is csatlakozik, külön kérdőívmintával. Ezt a felvételt egységesen 0,5 százalékosra tervezzük. Mintájához megyékben és megyei városokban a mikrocenzus lakásai közül minden negyedik kerül kijelölésre, mégpedig a jövedelmi felvétellel váltakozva, úgy, hogy e két felvétel közül egy lakásban csak az egyik kerüljön végrehajtásra. Budapesten a mikrocenzus számlálókörzeteit jelöljük ki váltakozva jövedelmi és átrétegződési felvétel céljára, ez utóbbiakban a mikrocenzus minden második lakásában kell az átrétegződési felvételi űrlapokat kitölteni.

6. Az intézeti háztartások mintája

Az intézeti háztartások mintájára az 1968. évi mikrocenzusnál alkalmazott és jól bevált eljárást irányozzuk elő.

A kiválasztási művelet előfeltétele, hogy időszerű címjegyzék álljon rendelkezésre. Szükséges tehát az 1970. évi népszámlálásból adódó címjegyzék kiegészítése és naprakész állapotba hozása. Az intézeteket megyénként, intézeti típusonként és a befogadóképesség alapján nagyság szerint csoportosítjuk.

Az 1968. évi eljárás szerint a különböző nagyságú intézetekből 10–100 százalékos arányban vettünk mintát és ennek megfelelően végezték a felvételnél az összeírandó személyek kiválasztását úgy, hogy végeredményben a mikrocenzus kívánt 2 százalékos aránya adódjon.

A 200 lakónál népesebb (illetve nagyobb befogadóképességű) intézeti háztartások valamennyien mintába kerülnek, az ennél kisebbeknél a táblázat szerinti arányban kiválasztást végzünk úgy, hogy lehetőleg az egyes megyék és a különböző típusú intézetek arányosan legyenek képviselve.

A személyek kiválasztása az előírányzott arányokban az összeírás helyszíni lebonyolítása alkalmával, a számlálóbiztosok útján történik. Ehhez névjegyzékre van szükség az intézeti lakókról, amelyet az intézet készített el, s ezen a számlálóbiztos a megadott kulcsszámok szerint az összeírandó személyek kijelölését végzi.

A személyek kijelölése többféleképpen történhet: az 1973 januári összeírás alkalmával, váltószámmal; a névjegyzék-úrlapon megfelelő sorok sorszámainak előre való bekarikázásával; felmerülhet kijelölt vezetéknev-kezdőbetűjű személyek, vagy bizonyos naptári napokon születettek mikrocenzusba vétele is.

A jövedelmi felvételt csak munkásszállásokon (esetleg főiskolai diákokthonokban) kell végrehajtani és csak azokról az intézeti lakókról, akiknek nincs családjuk, illetve háztartásuk, amelyhez tartoznának.

Az intézeti háztartásokkal kapcsolatban merül fel a jelenlevő és a lakónépesség összeírásának kérdése. A mikrocenzus adatainak feldolgozását lényegileg a lakónépességre alapozva tervezik; ilyen szempontból a kórházak, szállodák, üdülők összeírása a mikrocenzusnál mellőzhető is volna. Aggályos volna azonban a jelen-

levő népességszám ismeretéről lemondani, minthogy a lakónépességhez tartozó távollevők összeírása mindig kevésbé teljes, mint a jelenlevőké és így az új népességi fogalomra való áttérés is bizonytalan alapokon nyugodnék. Ezért egyszerűbb eljárással a jelzett intézeti háztartásoknál is veszünk 2 százalékos mintát.

7. Az új építkezések mintábavétele

Az első feladat a szükséges forrásadatok beszerzése. A népszámlálástól számított 3 esztendőről az adatok különböző időpontokban szerezhetők meg, mégpedig:

az 1970–1971. évi új építkezésekről a lakásépítési jelentések községi lajstromai vehetők igénybe; 1972. első feléről a megyei tanácsoknál levő használati engedélyekről készíthetők jegyzékek a Központi Statisztikai Hivatal megyei igazgatóságai útján.

1972 második felében folyó építkezések csak a mikrocenzus előzetes helyszíni lejárata során vehetők figyelembe.

Hasonló időbeli ütemezésben szerezhetők be a megyei igazgatóságok útján a lebontásokról szóló megszűnési jelentések.

A fenti anyagok beszerzését megelőzőleg szükség van előzetes tájékozódásra a lakótelepek tekintetében. Itt ugyanis többnyire szükség lesz a mechanikus módszerrel végzett számlálókörzet-kiválasztás módosítására, sok esetben a számlálókörzetek határainak módosítására, illetve új számlálókörzetek képzésére. Gondoskodni kell arról, hogy a lakótelepeknél is a lakások 10 százaléka kerüljön be az általános célú mintába és 2 százaléka a mikrocenzus-felvételbe.

A jegyzékekben szereplő új építkezések címeit hozzá kell dolgozni az 1970. évi népszámlálásból képezett minta anyagához. Ez a művelet természetesen csak azokra a községekre korlátozódik, amelyek a mintában szerepelnek. (Az új építkezések nagy része azonban erre az 1000 községre esik.) E községekben a számlálókörzeti térképek segítségével kell beosztani az új épületeket és hozzáírni a számlálókörzeti címjegyzékhez.

Minthogy a kisebb községekben 2–2 számlálókörzet, az 5000 lakoson felüliekben minden tizedik számlálókörzet szerepel a mintában, csak ezekre eső új épületeket kell a már korábban elkészült címjegyzékekre felvenni. E jegyzékből történik a mikrocenzus 2 százalékos mintájának megfelelően minden ötödik lakás kijelölése.

A címjegyzék kiegészítésével egyidejűleg ellenőrizni kell, hogy a megye összes új építkezéseinek valóban 10 százaléka került-e az általános célú mintába. Eltérés esetén az új lakások kijelölését megfelelő módon korrigálni kell.

Az új építkezésekkel is kiegészített címjegyzékek alapján a számlálóbiztosok 1972. november hóban bejárják körzeteik területét, ellenőrzik és kiegészítik a címjegyzéket. E módosítások alapján a Központi Statisztikai Hivatal megyei igazgatóságain – központi szakértőink közreműködésével – átvizsgálják a jegyzékeket és szükség szerint módosítják a mikrocenzus és kapcsolt felvételei során összeírandó lakásokat.

II. AZ ÁLTALÁNOS CÉLÚ ÉS A MIKROCENZUS-MINTA KÜLÖNBÖZŐ MEGOLDÁSI MÓDJAINAK VIZSGÁLATA

Az alábbiak néhány alapelvből kiindulva áttekintik az általános célú minta, illetve ezzel szoros kapcsolatban a mikrocenzus mintája kiválasztásánál követhető lehetséges eljárásokat.

A szóba jöhető mintavételi eljárások vizsgálatánál az alábbi néhány alapelvből kell kiindulni:

1. az általános célú minta kiválasztása az 5000-nél kisebb lélekszámú községekben két lépcsőben, a nagyobb településeken egy lépcsőben történik; ennek megfelelően a mikrocenzus mintája három-, illetve kétlépcsős mintavételi eljárással kerül kiválasztásra;
2. az általános célú minta a népszámlálási körzeteknek kb. 10 százalékát ölelje fel;
3. a községekre vonatkozó egy-egy rétegen belül valamennyi lakásnak (háztartásnak) egyforma esélye legyen a mintába kerülésre, azaz a háztartásokra vonatkozó kiválasztási valószínűség konstans legyen;
4. a mintába került körzeteken belül a lakások (háztartások) egyenlő valószínűséggel kerüljenek kiválasztásra.

A. Mintavételi eljárások az 5000-nél kisebb lélekszámú községek esetén

A megvizsgált nyolcféle kiválasztási eljárás mindegyike eleget tesz a fenti feltételeknek. Feltételezzük továbbá, hogy az eddigi gyakorlatnak megfelelően a végső kiválasztási egységek a lakások lesznek. A tárgyalás egyszerűsítése érdekében vezessük be az alábbi jelöléseket.

Legyen a községek egy adott rétegén belül

- N – a községek száma,
 - n – a kiválasztandó községek száma,
 - L_i – az i -edik községben a lakások száma,
 - $L = \sum_{i=1}^N L_i$ – a rétegben az összes lakások száma,
 - l – a kiválasztandó lakások száma a rétegben,
 - K_i – az i -edik községben a népszámlálási körzetek száma,
 - $K = \sum_{i=1}^N K_i$ – a rétegben az összes körzetek száma,
 - k_i – a mintába került i -edik községben a kiválasztandó körzetek száma,
 - $k = \sum_{i=1}^n k_i$ – az összes kiválasztásra kerülő körzetek száma,
 - L_{ij} – az i -edik község j -edik körzetében a lakások száma
- $$L_i = \sum_{j=1}^{K_i} L_{ij}$$
- l_{ij} – a mintába került i -edik község j -edik mintakörzetéből kiválasztásra kerülő lakások száma,
 - P_i – az i -edik község kiválasztási valószínűsége,
 - $P_{j|i}$ – az i -edik község j -edik körzetének kiválasztási valószínűsége, feltételezve, hogy az i -edik község bekerült a mintába,
 - $P_{\nu|ji}$ – az i -edik község j -edik körzetén belül a ν -edik háztartásra vonatkozó kiválasztási valószínűség, feltéve, hogy a szóban forgó körzet bekerült a mintába,
 - $P_{ij\nu}$ – annak valószínűsége, hogy az i -edik község j -edik körzetének ν -edik lakása bekerüljön a mikrocenzus mintájába

$$P_{ij\nu} = P_i P_{j|i} P_{\nu|ji} \quad /1/$$

A 3. alapelvnek megfelelően $P_{ij\nu} = 0,02$

A lehetséges mintavételi eljárásokat két nagy csoportba sorolhatjuk aszerint, hogy a községek kiválasztása *nagysággal arányos* vagy *egyenlő valószínűséggel* történik.

1. Községek kiválasztása nagysággal arányos valószínűséggel, községekben belül 2 körzet kiválasztása, azaz

$$P_i = \lambda L_i \quad i = 1, 2, \dots, N; \quad k_j = 1, 2 \quad i = 1, 2, \dots, n \quad /2/$$

1. A fenti feltételeken kívül kikötjük, hogy minden kiválasztott körzeten belül 17 lakás kerüljön be a mikrocenzusi mintába, továbbá, hogy a rétegen belül kiválasztott összes lakások száma a réteg lakásszámának pontosan 2 százaléka legyen, azaz

$$l = 0,02 L \text{ és } l_{ij} = 17 \quad i = 1, \dots, n; \quad j = 1, 2 \quad /3/$$

A 4. alapelv értelmében ekkor $P_{j|i} = \frac{17}{L_{ij}}$, továbbá /2/ és /3/ alapján

$$l = 17 \cdot 2 \cdot n = 0,02 L, \text{ amiből } n = \frac{L}{1700}$$

Mivel $P_i = \lambda L_i$ és $\sum_{i=1}^N P_i = \lambda \sum_{i=1}^N L_i = \lambda L = n = \frac{L}{1700}$, ezért $\lambda = \frac{1}{1700}$, vagyis:

$$P_i = \frac{L_i}{1700}$$

Igy /1/-et felhasználva

$$P_{ijv} = 0,02 = \frac{L_i}{1700} \cdot P_{j|i} \frac{17}{L_{ij}}, \text{ amiből: } P_{j|i} = 2 \frac{L_{ij}}{L_i}$$

Azt kapjuk tehát, hogy ha körzeten belül konstans számú lakást akarunk kiválasztani, akkor a községekben belül a körzeteket is *nagysággal arányos valószínűséggel* kell kiválasztani.

A kiválasztott körzetek száma: $k = 2n = \frac{L}{850}$, ami csak akkor adja pontosan a körzetek 10 százalékát, ha a rétegen belül a körzetek átlagosan 85 lakást tartalmaznak, azaz:

$$\frac{L}{K} = 85$$

2. Az $l_{ij} = 17$ feltétel helyett itt csak azt kötjük ki, hogy minden kiválasztott körzetből azonos számú lakást veszünk bele a mintába, viszont megköveteljük, hogy a minta a réteg körzeteinek pontosan 10 százalékát tartalmazza, azaz

$$k = 0,1 K \text{ és } l_{ij} = c \quad i = 1, 2, \dots, n; \quad j = 1, 2 \quad /4/$$

E feltételekből: $l = 0,02 L = kc = 0,1 Kc$, így $c = 0,2 \frac{L}{K}$, vagyis körzetenként a réteg átlagos körzeti lakásszámának 20 százalékát kell kiválasztani. Továbbá /2/-ből:

$$\sum_{i=1}^N P_i = \lambda L = n = \frac{k}{2} = 0,05 K,$$

amiből:

$$\lambda = \frac{0,05K}{L}, \text{ így: } P_i = 0,05 K \frac{L_i}{L}$$

/4/-ből és a 4. alapelvből:

$$P_{v|ij} = \frac{c}{L_{ij}} = 0,2 \frac{L}{KL_{ij}}$$

Végül az /1/ összefüggés alapján

$$P_{ijv} = 0,02 = 0,05 K \frac{L_i}{L} P_{j|i} = 0,2 \frac{L}{KL_{ij}},$$

amiből:

$$P_{j|i} = 2 \frac{L_{ij}}{L_i}$$

A körzeteket tehát itt is nagysággal arányos valószínűséggel kell kiválasztani. Az eljárás biztosítja a körzetek 10 százalékának kiválasztását. A rétegen belül a körzetekből azonos számú lakás kerül a mintába, ez a szám azonban különböző rétegekben különböző lehet a rétegre jellemző átlagos körzeti lakásszámtól függően.

Megjegyzendő továbbá, hogy mivel $0,2 \frac{L}{K}$ általában nem egész, a kerekítés miatt szükségessé váló korrekció következtében egyes körzetekben nem c , hanem $c-1$ vagy $c+1$ lakást kell kiválasztani.

3. Az előző kiválasztási eljárásoktól eltérően itt feltesszük, hogy a körzeteket a községeken belül egyenlő valószínűséggel választjuk ki, a kiválasztandó községek száma pedig megegyezik az 1. eljárásnál kapott értékkel, azaz

$$P_{j|i} = \frac{2}{K_i} \quad i = 1, \dots, N; \quad j = 1, 2, \dots, K_i \quad \text{és} \quad n = \frac{L}{1700}$$

Ez utóbbi feltételből – mint már az 1. eljárásnál láttuk – következik: $P_i = L_i : 1700$, s így /1/-ből

$$P_{ijv} = 0,02 = \frac{L_i}{1700} \frac{2}{K_i} P_{v|ji}$$

amiből:

$$P_{v|ji} = 17 \frac{K_i}{L_i}$$

és

$$l_{ij} = P_{v|ji} L_{ij} = 17 K_i \frac{L_{ij}}{L} = 17 \left(L_{ij} : \frac{L_i}{K_i} \right)$$

Itt tehát a körzeten belül kiválasztandó lakások száma körzetről körzetre változik, olyan mértékben tér el az átlagos 17-től, amilyen mértékben a szóban forgó

körzet lakásszáma eltér a községre vonatkozó átlagos körzeti lakásszámtól. A mintába kerülő lakások számát: l -et nem lehet előre rögzíteni, attól fog függni, hogy éppen mely községek, illetve körzetek kerültek bele a mintába; l természetesen 0,02 L körül fog ingadozni.

4. Körzetek kiválasztása, mint 3.-nál, egyenlő valószínűséggel, a mintakörzetek pontosan 10 százalékát ölelje fel, azaz

$$P_{j|i} = \frac{2}{K_i} \quad i = 1, 2, \dots, N; j = 1, 2, \dots, K_i \text{ és } k = 0,1 K$$

Az utóbbi feltételből – mint már a 2. eljárásnál láttuk – következik:

$$P_i = 0,05 K \frac{L_i}{L}$$

s így /1/ alapján

$$P_{ijv} = 0,02 = 0,05 K \frac{L_i}{L} \frac{2}{K_i} P_{v|ij}$$

amiből:

$$P_{v|ij} = 0,2 \frac{L}{K} \left| \frac{L_i}{K_i} \right|$$

A körzeten belüli kiválasztási arány tehát olyan mértékben tér el a 20 százaléktól, amilyen mértékben a rétegen belüli átlagos körzetnagyság eltér a községi átlagos körzetnagyságtól. A körzeten belül kiválasztásra kerülő lakások száma:

$$l_{ij} = P_{v|ij} L_{ij} = 0,2 L_{ij} \frac{L}{K} \left| \frac{L_i}{K_i} \right|$$

körzetenként változó, ennek következtében a rétegen belüli mintaelemszám nem rögzíthető előre.

II. Községek kiválasztása egyenlő valószínűséggel, a minta a körzetek 10 százalékát ölelje fel, azaz

$$P_i = \frac{n}{N} \quad i = 1, 2, \dots, N; k = 0,1 K \quad /5/$$

Ha a községeket egyenlő valószínűséggel választjuk ki, akkor – mint látni fogjuk – nem lehet rögzíteni előre a kiválasztandó lakások számát, feltéve, hogy ragaszkodunk az egyenlő esélyű mintába kerülés elvéhez.

1. Kiválasztott községekben belül két körzet kiválasztása nagysággal arányos valószínűséggel, azaz

$$k_i = 2 \quad i = 1, 2, \dots, n \text{ és } P_{j|i} = 2 \frac{L_{ij}}{L_i}, \quad i = 1, \dots, n; j = 1, 2 \quad /6/$$

/5/-ből és /6/-ból $k = 2n = 0,1K$, azaz $n = 0,05K$, és $P_i = 0,05 K : N$.

/1/ alapján:

$$P_{ijv} = 0,02 = \frac{0,05 K}{N} 2 \frac{L_{ij}}{L_i} P_{v|ij}$$

amiből:

$$P_{y|ij} = 0,2 \frac{N}{K} \frac{L_i}{L_{ij}} \text{ és } l_{ij} = P_{y|ij} L_{ij} = 0,2 L_i : \frac{K}{N}$$

A körzetekből kiválasztandó lakások száma tehát ebben az esetben községen belül állandó – az egy községre jutó átlagos körzetszámmal módosított községi lakásszám 20 százaléka –, községenként változó, a réteg mintaelemszáma így attól függ, mely községek kerültek be a mintába.

2. Községenként átlagosan 2 körzet kiválasztása nagysággal arányos valószínűséggel, körzeteken belül egységesen a rétegbeli átlagos körzeti lakásszám 20 százalékának kiválasztása, azaz:

$$P_{j|i} = \lambda_j L_{ij} \quad i = 1, \dots, n; j = 1, \dots, K_j \quad n = 0,05 K \text{ és } l_{ij} = 0,2 \frac{L}{K}$$

$$A \sum_{j=1}^K P_{j|i} = k_i \text{ azonosságból } \lambda_j = \frac{k_i}{K_j}, \quad i = 1, \dots, n; j = 1, \dots, k_j$$

azaz:

$$P_{j|i} = k_i \frac{L_{ij}}{L_i}$$

/1/ alapján:

$$P_{ijy} = 0,02 = \frac{0,05 K}{N} k_i \frac{L_{ij}}{L_i} \frac{0,2 L}{K L_{ij}}$$

Ebből:

$$k_i = 2 N \frac{L_i}{L} \quad /7/$$

A kiválasztandó körzetek száma tehát arányos lesz a község lakásszámban mért nagyságával. Itt látszólag előre rögzíteni lehet a mintaelemszámot, hiszen körzeteken belül a kiválasztandó lakások száma állandó és a $k = 0,1 K$ feltétel megszabja a körzetek számát is. Ez utóbbi egyenlőség azonban csak várható értékben teljesül. Ugyanis /7/-ből

$$k = \sum_{i=1}^n k_i = \frac{2 N}{L} \sum_{i=1}^n L_i$$

ami csak akkor adja meg a réteg körzeteinek 10 százalékát, ha a kiválasztott községek átlagos lakásszáma megegyezik a réteg átlagos községnagyságával, vagyis, ha

$$\frac{1}{n} \sum_{i=1}^n L_i = \frac{L}{N}$$

Ha ez a feltétel nem teljesül, akkor a minta által felölelt körzetek száma számottevően eltérhet a körzetek 10 százalékától, s így egyúttal a mintába került lakások száma is a réteg lakásainak 2 százalékától.

3. Községenként 2 körzet kiválasztása egyenlő valószínűséggel, azaz

$$k_i = 2 \quad i = 1, \dots, n \text{ és } P_{j|i} = \frac{2}{K_i} \quad i = 1, \dots, n; j = 1, 2 \quad /8/$$

/5/-ből és /8/-ből $n = 0,05 K$, így /1/ alapján

$$P_{ijv} = 0,02 = \frac{0,05 K}{N} \frac{2}{K_i} P_{v|ij}$$

amiből:

$$P_{v|ij} = 0,2 K_i \left/ \frac{K}{N} \right. \text{ és } l_{ij} = 0,2 L_{ij} K_i \left/ \frac{K}{N} \right.$$

A körzeten belüli kiválasztási arány tehát egy-egy községre nézve állandó és olyan mértékben tér el a 20 százaléktól, ahogyan a község körzeteinek száma eltér a rétegre vonatkozó átlagos községi körzetszámtól. A mintanagyság attól fog függni, hogy az első lépcsőben mely községek kerültek be a mintába.

4. Körzet kiválasztása egyenlő valószínűséggel, körzeteken belül egységes, 20 százalékos kiválasztási arány:

$$P_{j|i} = \frac{k_i}{K_i} \quad i = 1, 2, \dots, n; j = 1, 2, \dots, K_i \text{ és } P_{v|ij} = 0,2$$

Ezekből /1/ alapján:

$$P_{ijv} = 0,02 = \frac{n}{N} \frac{k_i}{K_i} 0,2$$

Ebből csak az következik, hogy a községekre, illetve a községen belül a körzetekre vonatkozó kiválasztási arányok szorzata 0,1 kell legyen. Ez többféleképpen is elérhető – például $0,4 \cdot 0,25$, vagy $0,2 \cdot 0,5$, vagy $0,3 \cdot \frac{1}{3}$, stb. –, a különböző lehetőségek közötti választás alapja lehet a községek közötti, illetve községeken belül a körzetek közötti átlagos szórások aránya, de lehet az is, hogy községenként átlagosan kb. két körzet kerüljön be a mintába. Bármilyen kiválasztási arányokat is rögzítünk, a mintába kerülő körzetek és háztartások tényleges száma a véletlentől függ, pontosan előre nem rögzíthető. Egy kétségtelen előnye viszont ennek a mintavételi eljárásnak, hogy a lakások számában időközben bekövetkező változások figyelembevétele csak a körzeti címjegyzék korszerűsítését igényli, a konstans kiválasztási arány következtében a változás arányosan tükröződni fog a mintában.

A fenti áttekintésből mindenekelőtt két következtetés vonható le. Először is a felsorolt alapelvek ugyan még többféle mintavételi eljárást tesznek lehetővé, de a különböző fokozatokban a kiválasztás módja mégsem lehet teljesen tetszőleges, illetve legfeljebb két lépcsőben választhatjuk meg a mintavétel módját. A másik következmény, hogy éppen a fent említett ok miatt a mikrocenzusi minta – s ezzel párhuzamosan az általános célú minta – kiválasztásánál alkalmazásra kerülő mintavételi mód lényegében megszabja az általános célú mintából a későbbi felvételekhez szükséges minták kiválasztási módját. Ha a mikrocenzushoz a kiválasztott körzeteken belül rögzített számú háztartást választunk ki, akkor a későbbi mintavételeknél is így kell eljárni, ha pedig adott kiválasztási arányt alkalmazunk, akkor a mikrocenzushoz hasonlóan a későbbi minták elemszáma sem rögzíthető előre pontosan.

Ami mármost az előbbieken felsorolt lehetőségek értékelését illeti, több szempont is az elsőként tárgyalt két *mintavételi eljárás* valamelyike mellett szól. Bár ezeknél a mintavétel előkészítése, lebonyolítása valamivel több munkát igényel, ugyanakkor szervezési szempontból kétségtelenül előnyös, hogy a kérdőbiztosok terhelése egyenletes és a mintába kerülő háztartások száma rétegenként előre rögzíthető. További előnynek tekinthető, hogy a tapasztalat szerint számos esetben a nagysággal arányos valószínűséggel történő kiválasztással (n. a. v. k.) kapott minta pontosabb, mint az ugyanolyan nagyságú, de egyenlő valószínűséggel kiválasztott (e.v.k.) minta. A n. a. v.k. minta pontossági előnye az e. v. k. mintával szemben általában annál nagyobb, minél szorosabb a korreláció a vizsgált ismerv és a nagyobb mintavételi egységek (községek, körzetek) nagyságát jellemző mérőszám – jelen esetben a lakások száma – között. Ezenfelül a hibaszámításra alkalmazandó formulák is egyszerűbbek n. a. v. k. minta esetén, ennek megfelelően számításigényességük is kisebb.

A szóban forgó két mintavételi eljárás (I. 1. és I. 2.) elméleti szempontból egyenértékűnek tekinthető, így közöttük gyakorlati megfontolások alapján lehet választani. Amennyiben a körzeti átlagos lakásszámban a rétegek között nincs jelentős eltérés, akkor inkább az I. 2. eljárás ajánlható, ellenkező esetben viszont szervezési szempontok inkább az I. 1. eljárás választása mellett szólnak.

B. Mintavételi eljárások az 5000-nél nagyobb lélekszámú települések esetén

Az egységes háztartási mintával kapcsolatos előterjesztésben lefektetett elveknek megfelelően az általános célú minta s ezzel együtt a mikrocenzus mintája is kiterjed valamennyi 5000 feletti lélekszámú településre. Így a mintavétel itt egyszerűbb és a lehetséges kiválasztási eljárások köre is szűkebb. Lényegében kétféle mintavételi mód között lehet választani:

1. A körzetek egy adott rétegén belül a körzetek kiválasztása nagysággal arányos valószínűséggel, körzeteken belül azonos számú lakás kiválasztása.
2. Körzetek kiválasztása egyenlő valószínűséggel, körzeteken belül azonos kiválasztási arány alkalmazása.

Bár nem szükségszerű, de célszerű, ha az 5000-nél kisebb, illetve nagyobb településeknél alkalmazott mintavételi módszerek szinkronban vannak, azaz, ha az 5000-nél kisebb lélekszámú községek esetén n. a. v. k. mintát alkalmazunk, akkor az ennél nagyobb településeknél is ilyen legyen a kiválasztási eljárás és fordítva.

A VITA SORÁN ELHANGZOTT FŐBB SZEMPONTOK

A vitaindító beszámolóhoz dr. Andorka Rudolf, dr. Ay János, Baranyai István, dr. Horváth Róbert, dr. Miltényi Károly, dr. Rácz Albert, dr. Schmidt Ádám, Schnell Lászlóné, dr. Szabady Egon, Tekse Kálmán, dr. Theiss Ede, dr. Vincze István és mások szóltak hozzá. Az élénk vita felölelte egyrészt a metodika, a kiválasztás és a rétegzés kérdéseit, másrészt az általános célú minta felhasználási területeit és témaköreit, végül szervezeti megvalósítása ügyét.

A mintavételi metodika szempontjából a hozzászólók megvitatták a rétegzés hatásosságát, tekintettel arra, hogy a községek kétféle (nagyság és foglalkozás szerinti) rétegzése egyes kategóriák átlépésével járhat. Az egyik hozzászólás javasolta, hogy elsődleges egységként közvetlenül számlálókörzetek kiválasztása történjék. Óvatosságot ajánlottak a minta korrekciója tekintetében, bár annak elvégzése csak a rétegek arányos képviselete érdekében és mechanikus eljárással van

tervbe véve. Megvitatták a megyénként különböző kiválasztási arány módszerét, ami a megyei becsléseket kétségkívül javítja, viszont más ismérveknél, például Budapestre koncentrálódó egyes társadalmi rétegek megfigyelésénél hátrányos is lehet. Szóba került a vezetéknev kezdőbetűje szerinti kiválasztás, amely módszer azonban egyes (például lengyelországi) tapasztalatok szerint esetenként torzítást is eredményezhet. A becslés pontosságának vizsgálatára a több részre bontott minta eljárásával egyetértettek, szükségesnek tartva a nem mintavételből eredő hibák vizsgálatát is. Sajnálatosnak tekintendő, hogy a most folyó hivatali adatgyűjtéseknek csekély százaléka reprezentatív jellegű, ami azt jelenti, hogy még nem sikerült kifejleszteni általánosan használt módszerré.

Az általános célú minta felhasználása tekintetében utaltak a társadalmi statisztikai adatok iránti igény növekedésére. Nemcsak pillanatnyi keresztmetszetekre van szükség, hanem az áramlási folyamatok nyomon kísérésére. Ilyenek például a különböző képzettséggel életbe lépők foglalkozási elhelyezkedése problémája stb. (szocio-demográfiai és szocio-ökonómiai vizsgálatok). Franciaországban a szak-képzettség növekedésének hatását vizsgálták meg társadalmi indikátorok útján, az oktatásügyi beruházásokkal összefüggésben. Longitudinális vizsgálatokhoz szükséges ugyanazoknak az egységeknek hosszabb ideig való megfigyelése. Össze kell azonban egyeztetni a longitudinális vizsgálatok iránti igényt a rotáció követelményével, mert a háztartások csak bizonyos ideig terhelhetők adatszolgáltatással.

A statisztika dinamikus jellegű feladatai mellett utaltak a háztartásra, mint egységre, amely nemcsak a fogyasztás szempontjából alapvető fontosságú, hanem szükséges a demográfia részleges átállítása is háztartási egységekre.

Mint további javaslatokat vetették fel a jövedelmi statisztika kiegészítését a vagyoni állapotról vonatkozó információkkal, továbbá a foglalkozási átrétegződésen kívül a képzettség változásának, a későbbi években való átképzésnek a felvételét. Az adott válasz szerint e javaslatok későbbi időpontban feltétlenül figyelmet érdemelnek.

Kérdés merült fel a mintának a háztartásoktól különböző egységekre, mint például mezőgazdasági üzemekre való alkalmazhatóságáról. Felvetették továbbá, hogy az általános célú minta a statisztikai hivatalon kívüli szervek részére is felhasználható lesz-e, „reprezentatív adatbank” jellegével.

Döntő feladatként jelölték meg a szervezet kialakításának és a számlálóbiztosok jó munkájának ügyét, amihez megfelelő személyek kiválasztása szükséges. Utaltak a kikerdezői munka nehézségeire, a kényes kérdésekre, ami a számlálóbiztosok megfelelő díjazását követeli meg. Nehézséget jelent, hogy a mikrocenzusnál a tervezett általános célú minta állandó jellegű kérdőbiztosi szervezete még nem fog fennállni. Minthogy azonban kisebb községekben csak egy-egy számlálóbiztos szükséges, remélhető, hogy a kétségkívül nehéz feladat megoldására lehet megfelelő embert találni.

A felvetett kérdésekre *Huszár István* államtitkár, a Központi Statisztikai Hivatal elnöke és az előadók válaszoltak. Az elnök kilátásba helyezte, hogy a módszertani szempontból lényeges javaslatokat megvizsgálják. A jövőben megvalósítandó folyamatos adatgyűjtések mintájának kérdését egyébként sem tartja lezártnak, a felmerülő kérdéseket állandóan vizsgálni fogják és szükség esetén a módszereket még munkaközben is módosítják.

A válaszban kilátásba helyezett vizsgálatot két megyében végezte el a Népszámlálás illetékes osztálya, még pedig arra a java latra vonatkozólag, amely községek kiválasztása nélkül, közvetlenül számlálókörzetek mintába vételét vetette fel, azok foglalkozási rétegzése szerint. A javasolt eljárással képzett minta, minthogy

kiválasztása szigorúan rétegek szerint haladt, szükségszerűen jól közelíti rétegenként a számlálókörzetek 10 százalékát, illetőleg a lakások 2 százalékát. A községkiválasztással induló, eredeti módszernél mutatkozó eltérések szintén nem nagymértékűek. Győr-Sopron megyében például 5 községben egy-egy számlálókörzet mechanikus cseréjével és két körzet pótlólagos behelyezésével a rétegek szerinti helyes arány biztosítható.

A javasolt közvetlen számlálókörzet-kiválasztás azonban erősen, kb. 50 százalékkal növeli a mintába kerülő községek számát. A kiválasztott községek túlnyomó részébe így módon csak egy számlálókörzet esik, ami vagy kevés munkát jelent, vagy két község közötti utazást tesz szükségessé. Az előirányzott 1000 község helyett ilyen módszer mellett kb. 1500 helyen kellene az összeírást végezni, ami szervezési és költségkihatások miatt nem valósítható meg. A vizsgált főbb demográfiai és foglalkozási ismérvek szempontjából, a teljes körű adatokkal összehasonlítva, mindkét kiválasztási eljárás egyaránt megfelelőnek mutatkozott.

Mindezek alapján az eredeti mintavételi terv mellett kell maradni, de a kérdés feivetése a mintaképzés több oldalú megvizsgálása és a tisztánlátás szempontjából tanulságos volt.

РЕЗЮМЕ

Настоящий отчет о состоявшемся 27 апреля 1972 года совместном заседании Статистической и Демографической комиссий Венгерской Академии наук останавливается на двух крупных задачах, которые в ближайшем будущем предстоит решить в области проведения выборочных обследований. Одной из них является намеченный на 1973 год микроцenz вместе с примыкающим к нему обследованием рестратификации населения по доходам и занятиям. Вторая задача заключается в создании такой выборочной совокупности, которая обеспечила бы возможность получения различных выборочных собраний данных по отдельным вопросам. Для этой цели предполагается создать такую единую систему сбора данных, которая располагала бы постоянной организационной структурой и подготовленными счетчиками, и в период между двумя переписями населения также выборочной совокупностью, позволяющей осуществлять актуальные расчеты.

Создаваемая для этой цели „выборка общего назначения” состоит из переписных районов переписи населения и охватывает примерно 10% последних. Выборка имеет двухступенчатую районированную структуру, на первой ступени производится отбор сел, а на второй — переписных районов.

Выборка микроценза — подобно таковой в 1963 и 1968 годах — является двухпроцентной. Территориальные единицы совпадают с таковыми в выборке общего назначения, но в переписных районах, отобранных согласно 2%-ой совокупности, в выборке микроценза фигурирует примерно каждая пятая квартира. Поскольку согласно плану выборочного обследования села и переписные районы попадают в выборку с вероятностью пропорциональной их величине, на третьей ступени из каждого переписного района в выборку включается одинаковое число квартир.

В ходе отбора сел производится районирование по величине групп населения и доле сельскохозяйственного населения. Все села и города, насчитывающие более 5000 квартир, попали в выборку.

Для определения квартир служит регистр переписи населения 1970 года, который дополняется адресами новостроек и упраздненных жилищ.

Примыкающее к микроцензу обследование доходов как правило является 1%-ным, за исключением Будапешта и четырех самых больших областей, где применяется 0,5%-ная выборка. К микроцензу примыкает обследование рестратификации занятий на основании 0,5%-ной выборки.

Во второй части отчета автор с помощью математического аппарата исследует способы, применяемые в ходе отбора выборки общего назначения и, соответственно, выборки микроценза. Это исследование исходит из того основного принципа, чтобы вероятность отбора в отношении квартир была равной (постоянной) как в случае того или иного района сел, так и в случае фигурирующих в выборке переписных районов.

Имея перед глазами этот основной принцип был осуществлен математический анализ восьми выборочных способов с целью выяснения вопроса о том, следует ли производить отбор сел и затем отбор переписных районов с вероятностью пропорциональной их величине или равной вероятностью.

SUMMARY

The report on the joint debate of Statistical and Demographic Committees of the Hungarian Academy of Sciences of 27th April, 1972 discusses two great sample surveys to be carried out in the near future. One of them is the micro-census intended for 1st January, 1973 along with related surveys on income and re-stratification of the population. The other one is to create such a sampling frame which can serve for different representative collections of data to be carried out in the forthcoming years. It is expected to form, with that aim, a uniform system of data collection with a permanent organization and trained enumerators, as well as a sampling frame to be brought up-to-date from one census to the other.

The „all purpose sample” to be formed consists of enumeration areas and covers some 10 per cent of them. It is built up by a twostep stratified sample, selecting communes in the first step and enumeration areas within them in the second.

The sample of the micro-census – similarly to the preceding ones of the years 1963 and 1968 – is of 2 per cent size. Its regional units are the same as in the all purpose sample, however within the enumeration areas selected in accordance with the 2 per cent representation, roughly every fifth dwelling is included in the sample of the micro-census. As according to the sampling schedule communes and enumeration areas fall in the sample with probabilities proportional to their size the same number of dwellings must be selected from every enumeration area in the third step.

The communes are stratified for sampling by population size-groups and by the share of agricultural population. All the communes and towns with more than 5,000 inhabitants will be sampled.

For the selection of dwellings the register of the 1970 census is used, completed with addresses of new buildings and demolitions that have occurred since then.

The income survey connected with the microcensus is usually of 1 per cent size except Budapest and the four largest counties in which a sample of 0,5 per cent is used. A survey of re-stratification of occupations, completed uniformly by 0,5 per cent sample, is also connected with the micro-census.

The second part of the report investigates the procedures that can be used for selecting the all purpose sample or that of the micro-census using mathematical methods. This investigation starts from the principle that the probability of selection of dwellings must be constant in each stratum of a commune, as well as in the enumeration area fallen into the sample. With this in view eight sampling methods have been investigated mathematically, depending on whether the selection of communes and that of enumeration areas is carried out with probability proportional or equal to their size.