

A REPREZENTATÍV LAKOSSÁGI ADATFELVÉTELEKRŐL

DR. MARTON ÁDÁM

Az elmúlt évtizedek során gyorsan növekedett a lakosság gazdasági és társadalmi helyzetének elemzése iránti igény, amihez következképpen egyre több információra, statisztikai adatra volt szükség. Ezt kielégítendő jött létre 1976-ban az Egységes Lakossági Adatfelvételi Rendszer (ELAR), amely az addig egymástól elszigetelten folyó, viszonylag kis számú reprezentatív lakossági adatgyűjtéseket szervezetileg és módszertanilag egységes keretbe foglalta, hosszú távú, egyeztetett programmal.¹

A kezdeti időszakot a módszertani útkeresés jellemezte. A tapasztalatok alapján 1983-ra alakult ki a jelenlegi formájában működő rendszer, az egységes mintavételi keret, amelyből a különböző felvételekhez azonos struktúrájú, bár sokszor különböző nagyságú mintát jelöltünk ki, valamint az egységes összeírói hálózat.

Az ELAR az egész országra kiterjed. Reprezentatív módszerrel „lefedi” az összes megyét, a különböző településtípusokat, módot ad a lakosság rétegeinek összehasonlító vizsgálatára. Vannak természetesen olyan felvételek is, amelyek csak a lakosság bizonyos részére (például az ifjúságra, a nyugdíjasokra stb.) vonatkoznak.

Az új ELAR működésének kezdete óta közel öt év telt el. Számos tapasztalat gyűlt össze a felvételek végrehajtása, elemzése, értékelése során. A különböző módszertani kérdések kezelése befolyásolja a végrehajtott felvételekből nyerhető információk, következtetések megbízhatóságát. Adódhatnak olyan problémák is, amelyek arra hívják fel a figyelmet, hogy a jövőben bizonyos kérdésekben (például a minta leterhelése) körültekintőbben, mértéktartóbban kell eljárni.

Az elmúlt évek munkája azt jelzi, hogy jó úton járunk. Az elmúlt időszak azonban arra is figyelmeztet, hogy érdemes számvetést készíteni, áttekinteni az elvégzett munkát s annak tapasztalatait hasznosítani a hátralevő öt év programjának előkészítésénél. Továbbá gondolnunk kell arra is, hogy milyen módon készítsük elő az 1993-ban kezdődő – feltehetően ismét tízéves – periódus programját.

Az ELAR keretében nagyon sok és változatos témájú felvétel végrehajtására került és kerül sor, amelyeknek szakmai problémáit, eredményeit számos elemző kiadvány és adatközlés ismerteti. A jelen tanulmánynak nem lehet célja az egyes felvételek szakmai sajátosságainak ismertetése. Csupán arra vállalkozhatunk, hogy vázoljuk a munkaprogramot, és áttekintsük a felvételek tervezésének, végrehajtásának és értékelésének általános módszertani kérdéseit.

¹ Az ELAR keretében folynak – néhány sajátos kivételtől eltekintve – a Központi Statisztikai Hivatal lakossági adatfelvételei. A rendszer működését részletesen ismerteti: Statisztikai Módszertani Füzetek 7. sz. Központi Statisztikai Hivatal, Budapest, 1984. 83 old.

A tanulmány két részből áll. Az első bemutatja az ELAR kereteit, programját, szervezetét. A második rész tárgyalja azokat a módszertani problémákat, amelyekre az eredmények felhasználásakor tekintettel kell lenni, illetve amelyeket a közeljövőben meg kell oldani.

I. AZ ELAR SZERVEZETE, PROGRAMJA

Az ELAR-felvételek közül a háztartási költségvetési felvétel (háztartásstatisztika) a legismertebb, amely az 1940-es évek vége óta folyik a Központi Statisztikai Hivatal keretei között. Jelenleg minden második (páratlan) évben van ilyen megfigyelés a mintavételi keret egyik részét alkotó H almintából kiválasztott 12 000 családdal közreműködésével.

A társadalmi statisztikai adatfelvételeket a páros években a másik, az ún. M almintán végezzük, amelyből minden évben 16 000 háztartást figyelünk meg. Ennek keretét a Társadalmi Demográfiai Alapadatok (TDA) felvétele (korábban a munkaerő-felvétel) adja meg, amely tartalmazza az alapvető demográfiai, társadalmi–gazdasági ismérveket (törzsadatokat). Ehhez kapcsolódik a többi felvétel, s a törzsadatok segítségével lehetőség van a különböző felvételek összekapcsolására, együttes elemzésére.

A felvételek fő célja az adott időszakra vonatkozó keresztmetszeti elemzések elvégzése. Az időben végbemenő változások vizsgálata a 2–5–10 évenként ismétlődő felvételek esetében a fontosabb társadalmi és demográfiai stb. csoportok szintjén történik.

A felvételek egy része rendszeresen, különböző periodicitással ismétlődik. Az említett háztartási költségvetési felvételre és a társadalmi–demográfiai alapadatok felvételére két évenként, a jövedelmi felvételre öt-, az időmérleg és a társadalmi mobilitás vizsgálatára tíz évenként kerül sor.

Az egyszeri vagy nem rendszeresen ismétlődő felvételek sok különböző témára terjednek ki: élet- és munkakörülmények, lakáshelyzet, művelődési szokások, foglalkozások presztízse, egészségi állapot, a nők társadalmi helyzete, nem nyilván tartható jövedelmek (második gazdaság), utazási szokások, elégedettség stb.

A minta

A legfontosabb gazdasági–társadalmi csoportosítások országos szintű reprezentációját biztosító minta kiválasztása három lépcsőben történt. Az első lépcsőben a település, a másodikban a körzet, a harmadikban pedig a lakás került kijelölésre.²

Az ELAR-minta településtípusok, valamint az eltérő sajátosságú társadalmi rétegek tekintetében nem arányosan rétegzett valószínűségi minta.

A kiválasztott családok, háztartások elvileg három éven át maradnak a mintában, ez azonban nem tekinthető hagyományos értelemben vett panel mintának, mivel az elemi szinten végbemenő változásokat mind ez ideig nem vizsgáltuk.

A minta a körzetek szintjén két részmintára oszlik: a háztartási költségvetési felvétel H mintájára és a többi társadalmi statisztikai felvétel M mintájára. A két minta teljesen azonos felépítésű, és minden olyan településre kiterjed, melyben legalább két körzet került kijelölésre. Az egy számlálókörzetből álló kis községek

² Részletes ismertetése megtalálható: Az ELAR-minta és az 1984. évi mikrocenzus mintájának kiválasztási eljárása. Statisztikai Módszertani Füzetek 24. sz. Központi Statisztikai Hivatal, Budapest, 1987. 59 old.

váltakozva kerültek a mintába. Végeredményben a H minta 382 településen 1316, az M minta pedig 390 településen 1314 körzetből áll.

A társadalmi statisztikai felvételek céljára az M mintából körzetenként 12 lakás, míg a háztartási költségvetési felvétel céljára a H minta körzeteiből 9 család került kiválasztásra. A kiválasztott címetek, illetve családokat a rotáció elveinek és gyakorlatának megfelelően egy, kettő vagy három éven át keressük fel, és kérjük arra, hogy közreműködésükkel segítsék elő munkánk sikerét.

Az ELAR-mintában az elsődleges mintavételi egység a település, illetve a körzet. Az összeírói munka egyszerűsítése, a költségek csökkentése érdekében szükséges, hogy „klaszterekből”, a meglehetősen koncentrált megoszlású körzetekből válasszuk ki a megfigyelésre kerülő egységeket, a családokat, illetve lakásokat. Mint a matematikai statisztika megfelelő megfontolásai alapján nyilvánvaló, azonos nagyságú (elemszámú) mintából, ha az teljesen véletlen kiválasztású, nagyobb pontossággal becsülhetők az egyes mutatószámok, mint ha a végső megfigyelési egységeket klaszterekből választjuk ki. Ez a megfontolás az ELAR-minta esetében azt jelenti, hogy azonos család- vagy lakásszám esetén minél nagyobb a körzetek száma, annál jobb eredményt kapunk.

Néhány felvétel céljára a becslések megbízhatóságának növelése érdekében mind a H, mind az M minta körzeteiből jelölünk ki háztartásokat. Ez az ún. J minta, amelynek terjedelme megegyezik az M mintáéval: az ELAR-minta összes körzetéből egyenként 6 háztartás került bele, így összesen mintegy 12 000 háztartást tartalmaz.

1983 és 1987 között a kiválasztott családok, illetve címek cseréje, a rotáció a következőképpen alakult.

1983-ban a háztartási költségvetési felvételt a H mintán körzetenként 6 családnál hajtották végre (8000 család), 1985-ben a minta 9 családra bővült oly módon, hogy körzetenként átlagosan 4,5 új családot kellett kijelölni a mintegy 1,5 családos „természetes” fogyás és a 3 családdal történő bővítés miatt. 1987-ben az 1985. évi mintában szereplő családok egyharmada került cserére. Így a 3 év összesített háztartásstatisztikai mintájában mintegy 2000 olyan család van, mely mind a 3 évben teljesített adatszolgáltatást, és körülbelül 6800 olyan, akik két alkalommal. Mivel 1989-ben a mintának azt az egyharmadát cseréljük ki, amely 1983 óta „él”, így nem lesz olyan család, amely 3 alkalomnál többször vett volna részt a háztartásstatisztikai adatgyűjtésben.

A H minta a rotációs elvnek megfelelően tehát olyan „guruló” minta, amelynél a minta kétharmada két egymás után következő éven át, egyharmada pedig három egymás után következő alkalommal szolgáltat adatot. (A természetes csökkenés miatt a tényleges arányok ennél valamivel kisebbek.)

Az M minta esetében más a helyzet. 1983–1984-ben nem került sor rotációra, míg 1985-ben teljesen új címek kerültek kijelölésre az M körzetekből. Két éven át a munkaerő-felvétel mellett a teljes M mintán még két másik felvételre került sor: vélemények az egészségügyről és az ifjúság helyzete. Ugyanakkor a jövedelem- és a mobilitásfelvétel J mintája lefedte az M minta felét. Így 1983–1984-ben kb. 8000 családnál öt felvételt hajtottunk végre. (Az elemi szinten összekapcsolható felvételek adatbázisba szervezése most folyik.)

Az ELAR-felvételek rendszere tehát olyan, hogy nemcsak egy éven belül, hanem két, esetleg három éven át is vannak olyan felvételek, amelyeket kisebb vagy nagyobb részben azonos mintán veszünk fel. Ez lehetővé teszi a felvételek összekapcsolt, komplex elemzését, természetesen amennyiben azok témái összefüggésbe hozhatók.

Az ELAR programja

1983 és 1987 között három háztartási költségvetési felvételre (1983, 1985, 1987) került sor.

A háztartási költségvetési felvételhez esetenként kisebb, kiegészítő jellegű felvételek is kapcsolódtak, amelyek segítségével meghatározott kiadásokat az adott évben a szokásosnál részletesebben kérdeztünk. Ilyen jellegű felvételek készültek az egészségügyi kiadásokról (1983), a kulturális és az üdülési kiadásokról (1985). Az 1983. évi háztartási költségvetési felvételhez kapcsolódva került sor a „Vélemények az életszínvonalról” című felvételre, amely a lakosság háztartási felszereltségében, anyagi-megélhetési helyzetében bekövetkezett változások következtében kialakított véleményét volt hivatott vizsgálni az életszínvonal alakulását illetően.

1983-ban volt a szokásos ötévenkénti jövedelmi felvétel, amelynek az a célja, hogy egy évre visszamenően részletesen felmérjük a lakosság, valamint fontosabb rétegei különböző forrásokból (beleértve a kisgazdaságokat és az egyéb forrásokat is) származó jövedelmét, annak változását, a strukturális eltolódásokat stb.

1983-ban került sor a tízévenként ismétlődő társadalmimobilitás-felvételre is. A vizsgálat 34 571 személyre terjedt ki. A kérdések a megkérdezett foglalkozás- és iskola-történetére, a nagyszülők, a szülők és a családtagok iskolai végzettségére és foglalkozására vonatkoztak. A vizsgálat célja a mobilitási esélyek időbeli, térbeli és társadalmi csoportok közötti alakulásának feltérképezése volt. A mobilitás-felvétellel párhuzamosan, annak 8413 személyre kiterjedő almintáján foglalkozás-presztízs-vizsgálatot végeztünk. Ennek során a megkérdezetteknek kártyákra írt foglalkozásneveket kellett rangsorolniuk olyan szempontok szerint, mint: a foglalkozással elérhető jövedelem, a foglalkozás gyakorlásához szükséges tudás, a foglalkozás társadalmi hasznossága, a foglalkozás általános presztízse, tekintélye.

1984-ben felmérés készült a lakoságnak az egészségügyi ellátásról alkotott véleményéről. 20 950 főt kérdeztünk meg. A kérdések többek között a lakóhely infrastrukturális ellátottságára (csatornázás, ivóvízzel való ellátottság, egészségügyi intézményhálózat közelsége stb.), a kérdezett személy és családja egészségi állapotára, egészségügyi intézményhálózat igénybevételi szokásaira, az igénybevétel gyakoriságára, az ellátás színvonalára, a gyógyszerellátásra vonatkoztak.

A fiatalok munka- és életkörülményeit volt hivatva felmérni az 1984-ben végrehajtott felvétel, amely: „Az ifjúságpolitika tudományos megalapozását szolgáló” tárcaszintű kutatási főirányhoz csatlakozott. Az egyszázalékosra bővített mintából 26 682 15–35 éves korú fiatalot kérdeztünk meg. A felmérés fel kívánta tárni a különböző iskolai végzettségű, családi háttérű rétegek foglalkozási (tanulási), anyagi és lakásvizonyait, valamint életkörülményeik, életmódjuk sajátosságait stb.

1984-ben mikrocenzus is volt, és annak keretében sor került több társadalomstatistikai felvételre is. A mikrocenzus mintegy 200 000 főre terjedt ki (2 százalékos minta). A felvétel a népesség szűkebb értelemben vett demográfiai szerkezete, iskolázottsága, foglalkoztatottsága, a háztartások, családok összetétele, a lakásvizonyok mellett számos más kérdést is vizsgált, minek következtében néhány, a tervekben szereplő ELAR-felvétel feleslegessé vált. Ezek többek között: a népesség egészségi állapotának néhány jellemzője, a felsőfokú végzettség hasznosítása a mindennapos munkában, az ingázás és a munkabajárással kapcsolatos közlekedés időigénye és módja, a gyermekek napközben történő ellátásának, gondozásának formái. A mikrocenzus kiterjedt a nyugdíjasok jövedelmi, vagyoni helyzetének, mentális, egészségi állapotának regisztrálásán túl étkezési szokásaik, szociális körülményeik, szabadidős tevékenységük, társas kapcsolataik felvételére is.

1986-ban felmértük a társadalmi–demográfiai alapadatok felvételének min-tájában szereplő személyek egészségi állapotát. A kérdőív olyan kérdéseket tartal-mazott, amelyek megfigyelésére az intézményi statisztika nem terjedhet ki. A viz-sgálat a következő főbb témakörökre vonatkozott: egészségi panaszok; az egézségügyi intézményhálózat igénybevétele; az életmód–életkörülmények azon vonatkozásai, amelyek kihatással lehetnek az egészségi állapotra (táplálkozás, dohány-zás, gyógyszeresedés, sportolás, munkakörülmények, munkahelyi ártalmak stb.).

Ugyancsak 1986-ban került sor a nők helyzetének felmérésére az M minta 16 000 háztartásának 15 éves és idősebb nem nappali tagozaton tanuló nőtagjai közreműködésével. A családtervezésre, a gyermeknevelés körülményeire, a ház-tartáson (családon) belüli munkamegosztásra, a háztartás felszereltségére, az otthoni munkát megkönnyítő szolgáltatások igénybevételére, a munkakörülmények-re vonatkoztak a kérdések.

A 15–79 éves korú népességből mintegy 11 000 főre terjedt ki 1986-ban a napi időbeosztást tudakoló időmérleg-felvétel. A mintába került személyeket négy al-kaalommal kérdeztük meg napi tevékenységeikről, minden periódusban az élet-módot befolyásoló egy-egy kérdésblokkal kiegészítve. A felmérés így nemcsak a megkérdezett egész napi (24 órás) tevékenységeinek, ezek társas környezetének, helyszíneinek kronológikus kikérdezésére szorítkozik, hanem a jövedelmi viszonyok-ra, a kulturálódási, üdülési, táplálkozási szokásokra, a családi és egyéb társas kapcsolati rendszerekre, a mobilitásra, a háztartások felszereltségére stb. is.

1986-ban olyan felvétellel is sor került, amely a nem számba vehető lakossági jövedelmek feltérképezését célozta. A családi (egyéni) kifizetések oldaláról mértük fel azokat a másoknak jövedelmet eredményező kiadásokat, amilyenek például a borralaló, a csúszópénz, a hálapénz stb. Az ilyen típusú jövedelmek áramlása nem nyilvántartott, sokszor nem legális csatornákon történik, ezért azok nyomon köve-tése, számbavétele az intézményi statisztika hatókörén kívül reked.

Az 1983 óta végrehajtott, valamint az 1988–1992. évekre tervezett felvételek-ről az alábbi felsorolás ad áttekintést.

Az ELAR keretében végrehajtott, valamint tervezett felvételek, 1983–1992

A felvétel megnevezése	Minta
1983.	
Jövedelmi	J
Háztartási költségvetési	H*
Egészségügyi kiadások	H*
Munkaerő	M
Társadalmi mobilitás (14 évesnél idősebbek)	J
Presztízsvizsgálat (14–70 évesek)	J-nek H körzetekre eső része
1984.	
Vélemények az életszínvonalról	H*
Vélemények az egészségügyről	M
Munkaerő	M
Fiatalok (15–35 évesek) életkörülményei	1 százalékra bőví- tett J
Mikrocenzus	2 százalékos
a) Ingatlantulajdon, lakásváltoztatási terv, lakásátalakítás	
b) Nyugdíjasok, nyugdíjaskorúak életkörülményei	
c) Longitudinális halandósági vizsgálat	
1985.	
Háztartási költségvetési	H
Kulturális kiadások	H
Üdülési kiadások	H


1986.	Társadalmi–demográfiai alapadatok (TDA)	M
	Nők helyzete (15 éves és idősebbek)	M
	Egészségi állapot	M
	Nem nyilvántartható jövedelmet eredményező kiadások	M
	Időmérleg–életmód (15–79 évesek)	1/3 M
	Pedagógus időmérleg–életmód	1300 fő
	Diák időmérleg–életmód (középfiskolás)	3000 fő
1987.	Háztartási költségvetési	H
	Háztartások mezőgazdasági termelése	H
1988.	Társadalmi–demográfiai alapadatok (TDA) a munkaerő- és réteghelyzet vizsgálatával	M
	Jövedelmi	J
	Presztízsvizsgálat (14 éves és idősebb)	M
	Egészségi állapot	M
1989.	Háztartási költségvetési	H
	Egészségi állapot	H
	Kulturális kiadások	H
	Lakosság fogyasztási és szolgáltatási kiadásai	H
	Lakosság közlekedési kiadásai	H
1990.	Társadalmi–demográfiai alapadatok (TDA) a lakókörnyezet megfigyelésével	M
1991.	Háztartási költségvetési	H
	Üdülési kiadások	H
	Háztartások felszereltsége	H
	Mezőgazdasági kistermelés	H
	Lakosság közlekedési kiadásai	H
	Áruvásárlási körülmények	H
1992.	Társadalmi–demográfiai alapadatok (TDA)	M
	Morbiditás, balesetek	M

Megjegyzés. Az egyes minták terjedelme és kiválasztása:

H* – 8 000 háztartás: a H körzetekből 6 család,

H – 12 000 háztartás: a H körzetekből 9 család,

M – 16 000 cím: az M körzetekből 12 cím,

 – 16 000 háztartás: az M minta fele, valamint a H körzetekből kiválasztott 8000 család (1988-ban 12 000+12 000 háztartás)

Az ELAR szervezete

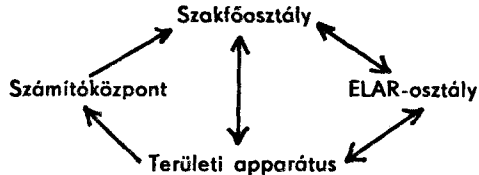
Az ELAR-felvételek kezdeményezése, szakmai irányítása a Központi Statisztikai Hivatal szakfőosztályainak feladata. A koordinációt és az általános módszertani fejlesztést a „központi” ELAR-osztály végzi. A területi apparátus felügyelete a Területi Statisztikai főosztály feladata. Az új minta kialakításával létrejöttek a felvételek ahhoz, hogy egységes kérdőbiztosi apparátus végezze az ELAR-adatfelvételeket. Valamennyi kérdőbiztos mind a háztartásstatisztikai megfigyelés, mind az egyéb lakossági témájú felvételek összeírásai munkáiban részt vesz.

A Hivatal megyei igazgatóságai mintegy 430 összeírót alkalmaztak 1985-ben. Az új minta keretei adta lehetőségeket kihasználva kedvezően alakult a főállásban foglalkoztatottak aránya, amely az összes összeírónak körülbelül 86 százalékát teszi ki. Az összeírók közül a minta „terítése” miatt elég sokan, (71 százalék, 303 fő) részmunkaidőben látják el a feladatot, mivel egy-egy településről gyakran nem annyi számlálókörzet került a mintába, amennyi egy vagy több teljes munkaidős összeíró alkalmazását tette volna lehetővé.

Az összeírók többsége, 81 százaléka egy településen, 15 százaléka kettő, 4 százaléka három településen végzi el az összeírási feladatokat. Az összeírók túlnyomó része 6–8 körzet összeírási munkáiért felelős.

Az összeírók munkáját közvetlenül irányító és az igazgatóságokon működő összeírási osztályok személyi állománya hozzávetőlegesen 120 fő, többségében közép- és felsőfokú iskolai végzettségű.

Az ELAR-szervezet működését némileg leegyszerűsítve a következő ábra szemlélteti:



A következőkben az ELAR elmúlt négyéves működésének tapasztalataiból adódó – részben megoldásra váró – problémákat vázoljuk fel.

Az ELAR-felvételek feldolgozása

Az eredeti koncepció szerint 1983-tól a felvételek kézi ellenőrzése, kódolása, hibajavítása átkerül a megyei igazgatóságokra, de a gépi ellenőrzés, feldolgozás, a táblázás, a különböző számítások, így például a hibaszámítás a Központi Statisztikai Hivatal Számítóközpontjában történik. A feldolgozás első fázisainak decentralizálása a hibajegyzékek segítségével a megyei igazgatóságokat segíti a kérdőívek, kódívek helyes kitöltésében. A folyamatos hibajavítás számos olyan tapasztalatot ad, amelyek felhasználhatók a soron következő felvételek előkészítésénél.

Az ELAR egyik rendszeres felvételének, a TDA-nak, korábban a munkaerő-felvételnek feldolgozása az előbb ismertetett menet szerint történik. A kérdőívek kézi ellenőrzését, kódolását a megyei igazgatóságok végzik. A központi adatrögzítés után a Számítóközpont ellenőrző programok segítségével vizsgálja az adatok logikai helyességét formai és tartalmi szempontból. A hibás tételekről a megyék listát kapnak, amit javítás után vissza kell küldeniök a Számítóközpontba, ahol a javított tételeket rögzítés után ismét ellenőrzik. A megyék általában két fordulóban kapnak hibalistát. Amennyiben már csak néhány, bizonylat nélkül is javítható hibás tétel marad, azokat a Számítóközpont az illetékes szakfőosztály segítségével javítja, majd elkészíti a jó adatokat tartalmazó állományt.

A többi társadalomstatisztikai felvételnél igen eltérő gyakorlat alakult ki az adatelőkészítés menetében.

Egyes felvételek esetén mind az ellenőrzést, mind a kódolást a szakfőosztály végezte (például a nők helyzete, 1986), más felvételeknél ezeket a műveleteket a megyei igazgatóságok szakmáymunkában készítették (például a nem nyilvántartható jövedelmet eredményező kiadások, 1986).

Több felvételre volt jellemző az a gyakorlat, hogy a kérdőívek zárt kérdéseit (az előre megadott válaszkategóriák alapján) az összeírók kódolták, a bonyolultabb, nyitott kérdéskörök (például táplálkozási szokások, olvasási szokások, betegségek) kódolása pedig központilag, az illetékes szakfőosztály irányításával történt.

Az adatok rögzítése, a gépi ellenőrzés és a hibalisták javítása – az említett munkaerő-felvételt kivéve – minden esetben központilag történt.

Az ELAR másik rendszeres felvételének, a háztartási és költségvetési felvétel adatainak feldolgozása 1987-ig központilag történt. Az 1987. évi felvétel adatainak rögzítése, kódolása és gépi ellenőrzése már a megyei igazgatóságok saját számítógépein történik. E felvételnél tehát az adatelőkészítés teljes decentralizálásáról beszélhetünk. A Számítóközpontba javítás után a megyei éves végleges állományok érkeznek majd a táblázási programok futtatására.

Az elsődleges feldolgozás valamennyi ELAR-felvételnél a hagyományosan szervezett állományokból egyedi programok használatával történt. A nem rendszeres felvételek esetén először típus táblákat kér a szakfőosztály, majd a különböző ismérvek gyakoriságának ismeretében kerül sor elemzési táblák készítésére.

Adatbázisban jelenleg csak az 1984-es mikrocenzusnak a személyre, a családra, a háztartásra vonatkozó adatai találhatóak.

A közeljövőben készül el – már meglevő rendszertervek alapján – másodlagos feldolgozások céljára az 1983-as jövedelmi felvétel, mobilitás- és presztízsvizsgálat és az 1984-es ifjúsági felvétel adatbázisa. Ezáltal lehetőség nyílik a felvételek adatainak együttes értékelésére. Folyik más felvételek adatbázisba szervezése is.

A kikérdezések megghiúsulásának alakulása

Az egyes felvételeket sohasem lehet – különböző okok miatt – a kijelölt teljes mintán végrehajtani. Az egyes kikérdezések megghiúsulhatnak azért, mert a kijelölt címek némelyikén üres lakás van, a lakók huzamosan távol vannak, vagy a mintába került személyek, családok közül néhányan megtagadják a válaszadást stb.

A megtagadások alakulása több tényezőtől függ. Így:

- a felvétel témájától,
- a megterhelés mértékétől (a kikérdezés időigénye),
- az összeírók rátermettségétől, kapcsolatteremtő képességétől,
- a válaszolók egyéniségétől, közreműködő-készségétől,
- bizonyos külső vagy egyedi tényezőktől stb.

A háztartásstatisztika esetében a megtagadások számát a többi felvételtől eltérő módon kell értelmezni, mert egyrészt „pótcimezés” történik: megghiúsulás esetében mindaddig keresünk újabb családokat, míg a körzetenkénti 9 család közreműködése biztosítva nincs. Másrészt, és talán ez a fontosabb, a háztartásstatisztikai napló két hónapon át tartó vezetése jelentős munkával jár, amit sokan nem vállalnak. (A „pótcimezés” módszere meglehetősen vitatható, bizonyos mértékig a reprezentativitás torzulásához vezet. Más, gyakorlatilag is megvalósítható módszer azonban nem áll rendelkezésre.)

Az emberek általában szívesen beszélnek az egészségi állapotukról, vagy ha más, őket érintő, számukra fontos dolgokról kérdezik véleményüket. Viszont a presztízsvizsgálat esetében például jóval nagyobb volt a megtagadások aránya, mivel mind a téma, mind a felvétel módszere sokak számára idegen, ismeretlen volt.

Az 1986. évi program nagy megterhelést jelentett a mintában levő családoknak. Az összes felvételt az M mintán kellett végrehajtani, s ebből az következett, hogy a család létszámától és a véletlentől függően egyes családokat a felvételi periódus során 12–16-szor is fel kellett keresni. Ugyanakkor az egyes kikérdezések időtartama 30–90 perc között alakult.

Az elmúlt öt évben végrehajtott felvételek során – amint ez a következő táblából kitűnik – a megghiúsulások száma nem volt jelentős.

A meghíúsulások alakulása 1983 és 1986 között

A felvétel megnevezése (ideje)	Kijelölt minta- elemszám	Megtagadások			Egyéb meghíúsulások		
		aránya (százalék)					
		vidéken	Buda- pesten	össze- sen	vidéken	Buda- pesten	össze- sen
1983-ban							
Háztartási költségvetési	7 896*	14,7	50,4	19,8	12,4	44,8	17,0
Munkaerő	15 756**	1,0	4,3	1,5	7,7	7,9	7,7
Jövedelem	15 780**	1,6	5,8	2,2	7,1	7,9	7,2
Társadalmi mobilitás	34 571***	2,3	6,3	2,8	3,6	5,8	3,9
Presztízsvizsgálat	8 413***	10,4	11,4	10,5	4,9	7,0	5,1
1984-ben							
Munkaerő (január)	15 756**	1,8	7,9	2,7	.	.	.
Munkaerő (december)	15 756**	2,3	9,7	3,3	.	.	.
Fiatalok életkörülményei	26 682***	1,7	5,5	2,1	6,9	10,8	7,4
Vélemények az egészségügyről	20 950***	0,4	1,5	0,5	2,9	5,3	3,2
1985-ben							
Háztartási költségvetési	11 844*	13,3	39,4	17,0	14,3	43,1	18,4
1986-ban							
Társadalmi–demográfiai alap- adatok (január)	15 756**	1,7	5,0	2,2	1,2	4,5	1,7
Társadalmi–demográfiai alap- adatok (szeptember)	15 756**	3,1	8,7	3,9	.	.	.
Időmérleg (első periódus)	11 241***	1,6	3,2	1,8	3,9	6,1	4,2
Időmérleg (második periódus) . .	10 696***	1,4	1,9	1,4	2,2	8,6	3,0
Időmérleg (harmadik periódus)	10 366***	2,4	1,6	2,3	3,6	6,0	3,9
Időmérleg (negyedik periódus)	9 626***	1,8	1,0	1,6	2,4	5,0	2,8
Nők helyzete	16 085***	2,1	4,6	2,4	1,1	1,9	1,2
Egészségi állapot	39 750***	2,2	1,7	2,2	3,4	4,5	3,6
Nem nyilvántartható jövedelmet eredményező kiadások	13 728*	0,5	–	0,4	1,2	–	1,1

* Háztartások.

** Lakások.

*** Személyek.

Megjegyzés. A táblában található adatok arra a körre vonatkoznak, amelyen a felvételt végre kellett hajtani. Mivel évközben folyamatosan van lemorzsolódás, az év végén végrehajtott felvételeknél már eleve az induló állománynál néhány százalékkal kisebb mintával kell számolni. Általában ugyanis a későbbi időpontban azokat, akik korábban megtagadták a válaszadást, már nem kell felkeresni.

A meghíúsulások mértéke eltérő Budapesten, a vidéki városokban és a falvakban. Ennek természetes magyarázata az, hogy a nagyvárosokban élők életvitele különbözik a kisebb településeken élőkétől, személytelenebbek a kapcsolatok, ami kedvezően befolyásolja a válaszok pontosságát (például a hálapénzekről a falvakban, ahol mindenki mindenkit ismer, nem szívesen nyilatkoznak).

A tapasztalatok a rendszer meglehetősen stabilitását jelzik, bár az összeírók gyakran panaszkodnak a válaszolók vonakodása miatt, főként hosszú időt igénylő kikérdezéseknél.

Annak, hogy 1986-ban nem következett be nagyobb mértékű megtagadás, részben az is a magyarázata, hogy az ELAR-apparátus, az összeírók a felvételek végrehajtása során sajátos, a többszöri felkeresésből adódóan hosszabb kapcsolatba kerültek az adatszolgáltató családokkal.

Az itt említett tény fontos sajátossága az ELAR-nak, amellyel számolni kell, de amivel visszaélni nem szabad. A válaszolókban felgyülemlett feszültség, ingerült-

ség nem tör tömegesen felszínre. Feltehető, hogy sokan vannak olyanok, akik szívesen elzárkóznának a felvételben való részvételtől, de az összeíróval kialakult személyes kapcsolatra, ismeretségre tekintettel mégis türelemmel fogadják a látogatót és kérdéseit. Ezek válaszaiban azonban valószínűleg több bizonytalanságot tartalmaznak, mint azoké, akik nem fáradtak el, nem ingerültek, érdeklődéssel, szívesen vállalják a válaszoló szerepét.

Az ELAR gyakorlata nem ismeri a részleges meghiusulást: azaz azt az esetet, amikor a megkérdezett egy-két kérdésre nem válaszol. A rendszer úgy épül fel, hogy az ún. „kitöltött” kérdőívben, minden kérdésre kell válaszolni, legalább annyit, hogy „nem tudom”, „nem emlékszem”, „nincs” stb. Ez az esetek túlnyomó részében nem okoz különösebb gondot, de valójában, nem egy esetben torzítás forrása lehet. Kérdőívszerkesztési szempontból megkíséreltük ezt a helyzetet úgy megoldani, hogy eleve kódolt, lehetséges válasz legyen a „nincs válasz” is. (A meghiusulások kezelésének módszertani problémáira visszatérünk.)

II. AZ ELAR MŰKÖDÉSÉNEK EGYES KÉRDÉSEI

A társadalmi és gazdasági folyamatokat, jelenségeket, összefüggéseket elemzők érthetően a lehető legpontosabb információkat szeretnék kapni. A statisztikai adatok „pontossága”, „megbízhatósága” módszertani kérdés, amit bármilyen statisztikai adatgyűjtésről legyen szó, részletesen tárgyalnak a különböző tankönyvek. Az, hogy az ELAR-felvételekből kapott adatok, a levonható következtetések megbízhatósága, pontossága fokozott figyelmet érdemel, nem szorul magyarázatra. A vizsgálatok reprezentatív jellege, a válaszok önkéntessége, a feltett kérdések sajátosságai olyan tényezők, amelyeknek megfelelő elemzése után lehet csak a kapott eredményeket helyesen értelmezni.

Az ELAR-jellegű felvételek olyan információk megszerzését célozzák, amelyek – nagyon kevés kivételtől eltekintve – más forrásból nem állnak rendelkezésre, nem vizsgálhatók kötelezően előírt teljes körű vagy mintán alapuló adatgyűjtések segítségével. Csak a tízevenkénti népszámlálások és az azok között, „félidőben” végrehajtott mikrocenzusok kivételek, amikor kötelező adatszolgáltatás írható elő. E munkának tehát nem hibája, hanem természetes velejárója, hogy szinte minden más statisztikai adatgyűjtéshez viszonyítva nagyobb mértékű bizonytalansággal kell számolni.

A felvétel³ végrehajtásának eszköztárához alapvetően a következőkből áll:

- megfelelő terjedelmű és struktúrájú minta;
- a célnak megfelelő kérdőív (utasítás, kódív stb.);
- hatékony számítástechnikai háttér;
- célszerű becslési eljárások, amelyek egyben biztosítják, illetve magukba foglalják a felvétel végrehajtásának értékelését, azt, hogy valóban megfelelt-e azoknak a megbízhatósági követelményeknek, amelyeket célul tűztünk ki a felvétel megtervezésekor.

A kapott eredmények, a becslések megbízhatósága döntően három tényezőtől függ:

- a mintavételi hibától,
- a nem mintavételi (számbavételi) hibától,
- a becslési módszer esetleges torzításától.⁴

³ Az ELAR-felvételek kérdőíves kikérdezések, ún. „face to face” interjúk.

⁴ Az elenzések, a kapott információk felhasználása során még további hibaforrások is lehetségesek, amelyek például a választott modellek illeszkedéséből vagy bizonyos következtetések interpretációjából adódnak.

Ha a felvételeket jól készítjük elő, körültekintően hajtjuk végre, és megfelelő formulákkal számolunk, akkor „jó” eredményeket fogunk kapni. Miből áll a jó előkészítés, és hogyan mérhető (becsülhető) az eredmény jósága? Ezzel foglalkozunk a továbbiakban a teljesség igénye nélkül.⁵

A felvételek végrehajtásának és értékelésének eszközei

Az elmúlt öt év tapasztalatai arra utalnak, hogy a minta alapvetően megfelel az eredetileg támasztott követelményeknek. Bizonyos megfontolások azonban azt jelzik, hogy szükséges lenne az adott keretek között bizonyos kiegészítő, módosító jellegű változtatásokat végrehajtani.

A minta rögzített terjedelmű: az egyszer kiválasztott címek – mint arról már volt szó – a rotáció elveitől függően két-három éven át, egy-egy év kihagyással a mintában maradnak.

A kiválasztott családokat még akkor is, ha azok nem minden tagját kell a különböző felvételek során megkérdezni, évente több alkalommal kell felkeresni. Ha pedig egy-egy látogatás során a megkérdezés hosszú időt vesz igénybe (egy, esetleg másfél órát), számolni kell a válaszolók türelmének kimerülésével, ami szükségszerűen a megtagadások és (vagy) az át nem gondolt, megbízhatatlan válaszok számának növekedéséhez vezet. Erre a problémára a minta oldaláról csak úgy lehet megoldást találni, ha azt rugalmassá, változó terjedelművé tesszük oly módon, hogy az adott felvételi program több válaszoló között kerüljön megosztásra. A megfelelő tervezés biztosíthatja az összefüggő felvételek összekapcsolhatóságát is.

Az utóbbi években a szakirodalomban számos tanulmány jelent meg arról, hogy az ELAR-jellegű mintákat hogyan lehet időben és térben úgy megosztani, hogy azok külön-külön is értékelhető eredményeket adjanak, egyesítés után pedig egy magasabb szinten komplexebb elemzést tegyenek lehetővé.⁶

Célszerű lenne például olyan megoldást keresni, hogy az M mintát két egyenlő nagyságú M_1 , M_2 részre osztva, a tervezett felvételeket az összekapcsolás szempontjait is figyelembe véve csak valamelyik részmintán hajtjuk végre. Az M minta megosztása a címek szintjén történhetne meg, s így mindkét rész minta tartalmazná az összes körzetet. A felvételek jellegétől függő mintanagyságot, amennyiben az M_1 vagy az M_2 minta nem lenne elégséges, úgy kellene biztosítani, hogy további címeket jelölünk ki az M minta körzeteiből. E címeket, hogy az 1992-ig rendelkezésre álló címanyag ne merüljön ki idő előtt, a későbbi rotáció során figyelmen kívül kellene hagyni. Így minden felvétel mintája az M_1 (vagy M_2) állandó, valamint az ahhoz szükség szerint hozzárendelt változó részből állna. E változó rész megtervezése megynként történhetne, figyelembe véve ezáltal a megyei tájékoztatási igényeket is.⁷

Így ki lehetne alakítani a minták olyan rendszerét, amelyben az egyes minták között részleges átfedés lenne (esetleg időbeni megosztottsággal). A kérdőívek egyes blokkjainak kellő megtervezésével biztosítani lehetne a különböző kérdéskörök megfelelő megbízhatóságú (megfelelően nagy elemszámú mintán ala-

⁵ Speciális szakmai módszertani problémák (mint például az osztályozás vagy a vizsgálni kívánt mutatószám-rendszer célszerűsége, amelyek az elemzések sikeressége szempontjából nagyon fontosak) nem képezik vizsgálódásaink tárgyát.

⁶ Lásd például Kish, L.: Samples and censuses. *International Statistical Review*. 1979. évi 47. sz. 99–109. old.; Kish, L.: A közérdekű döntések előkészítését szolgáló felvételek időzítése. *Statistikai Szemle*. 1985. évi 1. sz. 5–14. old.

⁷ A minta nagyságának meghatározásakor azt is figyelembe kell venni, hogy milyen területi részletességű információkra van szükség. Megyei eltérések vizsgálatához természetesen nagyobb mintára van szükség, mint országos szintű becslések céljára.

puló) becslését, valamint a megyei reprezentativitás növelését. A probléma kidolgozása során tekintettel kell lenni a minta nem arányos voltából adódó kiválasztási és teljeskörűsítési problémákra.⁸

A jelenleg működő ELAR-minta egyik legvitatottabb jellemzője a nem arányos volta. Az a cél, hogy a homogénebb társadalmi rétegek közül viszonylag kevesebb családot válasszunk ki, vitathatatlanul helyes, gazdaságos. A gyakorlatban azonban ez az elv nem mindig, illetve minden szempontból valósítható meg. Nincsenek „homogén” körzetek, és a mintavétel időpontjában nem állnak rendelkezésre, illetve csak igen nagy munka árán lennének elérhetőek megfelelő információk a körzetek lakosságáról.

A minta a településnagyság és a társadalmi rétegződés szempontjából nem arányos.

Budapest részesedése a mintában kb. 14 százalék a tényleges 21 százalékkal, a vidéki városoké 31 százalék a valóságos 33 százalékkal szemben, míg a községeké a ténylegesnél nagyobb, 55 százalék a népességszám szerinti 45 százalék helyett.⁹

A településnagyság szerinti aránytalanságot jelzi az is, hogy a különböző nagyságú települések körzeteinek hány százaléka került be a mintába.

A kijelölt körzetek aránya a mintában

Lakosok száma (fő)	Körzetek aránya (százalék)	Lakosok száma (fő)	Körzetek aránya (százalék)
Községek		Városok	
– 1 000	10,0	– 10 000	8,5
1 000– 1 499	7,3	10 000–14 999	7,0
1 500– 2 999	8,8	15 000–24 999	6,8
3 000– 4 999	8,4	25 000–	6,3
9 000– 9 999	7,0	Városok együtt	6,5
10 000–14 999	7,8	Megyei városok	5,5
15 000 és több	7,0	Budapest	4,6
Községek együtt	8,3	Ország összesen	6,8

A nem arányos kiválasztás másik szempontja és célja az volt, hogy a munkások, szellemiek és inaktívak a tényleges súlyuknál kisebb, míg a szövetkezeti parasztság és a kettős jövedelműek – minthogy az ilyen családok jövedelmi és egyéb viszonyai sokkal kevésbé homogének, mint az előbb említett rétegeké – nagyobb arányban legyenek reprezentálva. Mivel a körzetek népessége mindig több társadalmi rétegbe tartozik (nincsenek tisztán paraszti vagy munkás, illetve értelmiségi vagy nyugdíjas körzetek), e tekintetben a minta tervezett „aránytalanságát” megvalósítani nem sikerült.

Az ELAR-minta reprezentativitása ellenőrizhető a személyekre és a lakásokra vonatkozó legfontosabb adatok alapján, amelyeket minden felvételnél – mint ún. törzsadatokat – számba veszünk. Így kiszámítható például a népesség száma és megoszlása bizonyos területi és egyéb csoportok szerint, valamint néhány adat a lakásokra vonatkozóan is. Ezek összevethetők a népszámlálás teljes körű, továbbvezetett adataival, amelyek országosan majdnem pontosnak tekinthetők. Amennyi-

⁸ Külön lehetőség a minta megbízhatóságának javítására – adott elemszám mellett – a körzetek számának növelése. Az M körzetek címei mellett a H körzetek címeinek felhasználása további lehetőséget kínál az adott mintavételi keret rugalmas felhasználására. Ez történik például a jövedelemfelvételek J mintái esetében.

⁹ Az új minta indulásakor (1983. január 1-én) érvényes közigazgatási beosztásnak megfelelően.

ben a teljes körű és a számított adatok között 2–3 százalékos eltérésnél több mutatkozik, célszerű megvizsgálni mi okozta a különbséget, és el kell dönteni, szükséges-e valamelyes korrekció. (Az ún. intézeti háztartásokra az ELAR-minta nem terjed ki!)

Ilyen számítások jó kontrolljai a mintának, ami – különböző okok, így például a meghiúsulások miatt – felvételenként valamelyest mindig változik. Hangsúlyozni kell azonban, hogy ezek az összehasonlítások nem pótolják a hibaszámítást. Ha a törzsadatok összehasonlítása kedvező eredményt ad is, abból még nem következik, hogy a becsült mutatószámok hasonlóan közel vannak azok tényleges értékéhez.

Az ELAR keretében néha olyan felvételekre is sor kerül, amelyekben bizonyos viszonylag ritkán előforduló jelenségeket (például betegséget) is vizsgálunk. Az általános célú ELAR-mintától nem várható el, hogy az ország népességének csak kisebb részét érintő kérdéseket is jól reprezentálja. Emlékeztetni kell arra, hogy például a népesség kevesebb mint egyharmadát (14–35 évesek) kitevő ifjúság helyzetének vizsgálatok a mintát már célszerű volt százszázalékosra bővíteni.

Tehát az, hogy az ELAR-minta sem családi, sem személyi szinten nem reprezentál bizonyos jelenségeket, a minta kiválasztásából természetesen adódik. Speciális felvételeknél (például csökkent munkaképességűek vagy bizonyos betegségben szenvedők helyzetének vizsgálatánál) a mintát jelentősen meg kell növelni, vagy más módon, a működő ELAR-mintától független mintát kell kidolgozni.

Viszonylag rövid idő múlva sor kerül az új, 1992 utáni minta alapelveinek kidolgozására. Ehhez csupán annyit jegyzünk meg, hogy a mintavételi keret a jövőben is valószínűleg a népszámlálás címanyaga lesz. A minta nem arányos jellegén azonban célszerűnek látszik változtatni: a nem arányos kiválasztást csak a településnagyságtól tenni függővé, a jelenleginél esetleg nagyobb különbségeket megengedve.

A kérdőívek megszerkesztése hosszadalmas és meglehetősen kellemetlen munka. Azok, akik szerkesztettek már kérdőívet, tudják, hogy mennyi probléma adódik a lehetséges válaszok jó besorolásából, a kérdések sorrendjéből és logikai összefüggéseiből, illetve annak hiányából, a csoportosítások és osztályozások következetes, egységes alkalmazására való törekvésből stb. Néha még magukat a megjelölt adatokat, a háztartáson (családon) belüli viszonyokat sem könnyű egyértelműen definiálni, különösképpen akkor, ha egy paneljellegű mintán több felvételt is végrehajtottunk, s azok között a kapcsolódást is biztosítani akarjuk.

Az ELAR-felvételek kérdőíveinek előkészítése többoldalú egyeztetéssel történik. Hiányzik azonban annak az igazán „tudományos” megközelítése. A felvételek szakmai előkészítését végzők egy kicsit szükséges rossznak tekintik az arra fordított időt. Teljesen érthető, hogy számukra a téma vizsgálata a fontos, nem pedig a kérdőívszerkesztés elméleti kérdéseiben való elmélyülés. Az utóbbi hónapok során az ELAR új szervezeti keretei között történtek bizonyos lépések a korábban kialakult KSH-kérdőív-szabvány korszerűsítésére, s az 1986. évi felvételek végrehajtása már eszerint történt. Ez azonban csak szerény kezdet.¹⁰

A kérdőívszerkesztés elméleti problémáit a nemzetközi szakirodalom részletesen tárgyalja. Ebből az is megállapítható, hogy e téren sok országban vannak hiányosságok. A probléma aktualitását jelzi, hogy a svéd Statisztikai Hivatal folyóirata a közelmúltban egy egész számot szentelt e kérdések tanulmányozásának.¹¹

¹⁰ Lásd bővebben: A kérdőívtervezés néhány általános problémája. ELAR Módszertani Tanulmányok. 2. sz. (Kézirat.)

¹¹ *Journal of Official Statistics*. 1985. évi 2. szám. Special Issue on Questionnaire Design.

Ahhoz, hogy a kérdőíveknek, a kérdések megfogalmazásának és ezzel összefüggésben az összeíróknak szerepét tanulmányozzuk, kísérleteket kellene végezni

- alternatív kérdőívekkel,
- ismételt kikérdezésekkel (follow-up),
- rátermett és kevésbé rátermett összeírókkal stb.

Sajnos azonban ez idő szerint ez nem szerepel a tervünkben.

A becslési formulák megválasztása lényeges kérdés a felvételekből kapott eredmények megbízhatóságát illetően. Az utóbbi időben számos olyan dolgozat jelent meg, amelyek azt vizsgálták, hogy bizonyos pótlólagos információk (modellek) felhasználásával miként lehet a valószínűségi mintákon alapuló becslések megbízhatóságát növelni.¹²

Mint arról szó volt, az ELAR-minta célja országos szintű reprezentativitás biztosítása volt a legfontosabb területi és társadalmi rétegek szerint. Amennyiben olyan igény merül fel, hogy bizonyos mutatószámok társadalmi rétegenkénti eltéréseit megyei szinten is kimutassuk, akkor az ELAR-minta terjedelme általában már nem elég nagy. Ilyen esetekben – esetleg – segítséget nyújthatnak az ún. kisterületi becslések, melyekkel már több tanulmány foglalkozott, s így ezen a helyen ennek tárgyalásától eltekintünk.¹³

A felvétel eredményeinek megbízhatóságáról

A matematika: statisztika jól ismert fejezete a reprezentatív felvételekből származó becslések megbízhatósági határainak, a mintavételi hibának meghatározása. Még a bonyolult mintavételi tervek esetén sincs elvi akadálya a hibák becslésének. A hibaszámítás burkoltan azt feltételezi, hogy a mintában levő adatok pontosak, azaz nincs számbavételi (nem mintavételi) hiba.

A reprezentatív felvételekből, különösen az ELAR-felvételekből nagyon sok irányú információkat közölnek, amelyeknek megbízhatósága részletezettségük és a vizsgált mutató viselkedésének sajátosságai miatt eltérő. Természetes, hogy egy adott felvételtől nem minden változó becslhető egyformán jól.

A jelenlegi meglehetősen bonyolult felépítésű ELAR-minta esetében hasonlóan bonyolult a mintavételi hiba kiszámítása. A rendelkezésre álló számítástechnikai lehetőségek, a CLUSTERS programcsomag¹⁴ segítségével csak nagyon nehézkesen, hosszadalmasan lehet a mintavételi hibát kiszámítani. Ennek rendszeres és tömeges alkalmazására érthető okokból nem került sor.

1986-ban kifejlesztettünk egy új hibaszámítási programot, amelynek segítségével az előbb említett programcsomagnál egyszerűbben és gyorsabban lehet az ELAR-mintán végrehajtott felvételeknél hibát számítani.¹⁵ Ez a program ugyan valamilyen „kevesebbet tud”, mint a CLUSTERS, de úgy véljük, hogy előnyei messze meghaladják a bizonyos „szolgáltatások” hiányából adódó hátrányt. Alkalmazását megkezdjük, és kellő tapasztalatok birtokában sor kerülhet továbbfejlesztésére, például arra, hogy alkalmassá tegyük a megyei statisztikai igazgatóságok TPA-gépein történő felhasználásra is.

¹² Ezekről áttekintést ad: dr. Marton Ádám: A mintán alapuló becslések néhány terminológiai és módszertani kérdése. *Statisztikai Szemle*, 1987. évi 4. sz. 310–315. old.

¹³ E problémák részletes tárgyalását lásd a „Területi és egyéb szempontok szerint részletezett statisztikai mutatószámok becslése” (Központi Statisztikai Hivatal, Budapest, 1985. 146 old.) c. kiadványban.

¹⁴ CLUSTERS—A package program for computation of sampling errors for clustered samples. International Statistical Institute, World Fertility Survey, London, 1978. 73 old. A „CLUSTERS” ebben az esetben a Computation and Listing of Useful Statistics on Errors of Sampling megnevezés rövidítése.

¹⁵ Az új program használatát az „Új CLUSTERS — egy egyszerű lehetőség az ELAR mintavételi hibák kiszámítására” c. anyag ismerteti. ELAR Módszertani Tanulmányok. 3. sz. (Kézirat.)

A becslési módszerekkel kapcsolatban már szó volt arról, hogy vannak olyan speciális módszerek, amelyek bizonyos feltételek teljesülése esetén lehetővé teszik a kisterületi (megyei) szintű becslések javítását. Részben a megyei szintű becslési lehetőségek vizsgálata, részben a hibaszámítási módszerek tanulmányozása céljából több megyei szintű becslésre vonatkozó hibaszámítást végeztünk el.¹⁶

E munka során több gyakorlati, szemléletformáló tapasztalat adódott. Csak nagyon vázlatosan: a megyei szintű ELAR-becslések hibahatárai meglehetősen nagyok. A becslőt mutató természetétől, viselkedésétől függően 50–100 százalék körül is alakulhatnak. Különösen nagyok a relatív hibahatárok megoszlások esetén, amikor valamelyik „cella” részesezése kicsi, csak néhány százalékos. Arra tehát, hogy az egyszerű ELAR-becslések segítségével megyénkénti szignifikáns különbségeket ki tudjunk mutatni, csak összefoglaló mutatószámok esetében van valamilyen remény, amennyiben egyáltalán lényeges megyei eltérés létezik. Ez azonban mégsem jelenti azt, hogy a kapott információk megyei szinten semmire sem lennének használhatók.

Számos esetben a vizsgált jelenségek megoszlása igen jelentős eltéréseket mutat, például a családi konfliktusok forrásai között az ital szerepe 10–15-ször nagyobb, mint a lakásproblémáké. Ez az eltérés, tehát az ital nagy szerepe, megyénként külön-külön is jól megállapítható. Nem lehet viszont azt vizsgálni, hogy az italozás 10 százalékpont körüli részesezése megyénként összehasonlítva különbözök tekinthető-e, mivel 1–3 százalékpont eltérés nem tekinthető szignifikáns különbségnek.

Ez a megjegyzés arra az egyébként magától értetődő, de sokszor nem kellően értelmezett jelenségre hívja fel a figyelmet, hogy még az igen nagy hibahatárok között becsülhető mutatószámok is bizonyos szempontból értékelhetők. Az a lényeges kérdés, hogy mi a vizsgálat célja.

A megyei adatok iránt megnövekedett érdeklődésre tekintettel 1986-ban azt is megvizsgáltuk, milyen következményekkel járna a minta elemszámának növelése. Az ELAR-minta jelenlegi kereteit az anyagi lehetőségek befolyásolják. Felvetődött azonban a kérdés, hogy ha egy-egy felvétel esetében szükség lenne megyei szinten is megbízhatóbb adatokra, a mintát milyen mértékben kellene növelni ahhoz, hogy az eredmények érezhetően javuljanak.

A vizsgálat módszere az volt, hogy az 1984. évi mikrocenzus anyagából generáltunk különböző méretű (a jelenleginél nagyobb) mesterséges ELAR-mintát, s öt különböző nagyságú és jellegű megye esetében 10 különböző mutató becslését és hibaszámítását végeztük el. Az általános megfontolásokkal összhangban olyan eredményeket kaptunk, hogy a minta 10–20 százalékos növelése nem hoz jelentős eredményt. Az 50–100 százalékos növelése az, ami már a konfidencia intervallumok számottevő csökkenéséhez vezet.¹⁷

Ugyanez a kísérletsorozat azt is vizsgálta, hogy országosan miként alakul a becslések megbízhatósága a minta elemszámának csökkentése, a körzetek számának változatlanlansága, illetve növelése esetén. A felvételek végrehajtásával kapcsolatos munka túlnyomórészt a minta elemszámától függ. Az adott mintavételi kereten belül a körzetek számának növelése – például mind a H, mind az M körzetek felhasználása, amint az a J minta esetében történik – alig növeli a „terepmunkát”. A szimulációs kísérletek viszont azt jelezték, hogy a mintavételi hiba nem minden esetben csökkent. Ez az eredmény további vizsgálatot igényel.

¹⁶ E munka eredményeiről „A fiatalok életkörülményei” c. 1. sz. ELAR Módszertani Tanulmányban számoltunk be. (Kézirat.)

¹⁷ Lásd bővebben: Az ELAR-felvételekből származó adatok mintavételi hibájának alakulása a mintanagyság lehetséges csökkentése, illetve növelése esetén. ELAR Módszertani Tanulmányok. 4. sz. (Kézirat.)

A mintavételi hiba nagysága több tényező mellett a mintavételi tervtől is függ. Viszonylag új fejezete a mintavételi hiba vizsgálatának komplex, sokváltozós és sokféle bontást (társadalmi réteg, terület stb.) igénylő becsléseknél az, hogy nagyságát miként befolyásolja a praktikus megfontolásokat tartalmazó, tehát a véletlen hatását korlátozó mintavételi terv. Erre szolgál az ún. „roh” és „deft” kiszámítása.¹⁸ E kérdés vizsgálata folyamatban van, részletezése azonban meghaladja e tanulmány kereteit.

A statisztikai adatok pontosságát befolyásoló másik tényező a számbavételi, a nem mintavételi hiba. A teljes körű, kötelező vállalati, intézményi adatgyűjtések esetében is előfordulhat, hogy a jelentett adat valamilyen okból kifolyólag hibás. Ebben a körben azonban lehetőség van az adatok ellenőrzésére, ami sok esetben meg is történik.

A lakossági, önkéntes adatfelvételek körében a számbavételi hiba – a válaszadási hatás – a kérdések jellegétől függően eltérő mértékben a kötelező intézményi adatgyűjtéseknél sokkal nagyobb.

Röviden mi e probléma lényege? A személyekhez intézhető kérdések alapvetően két csoportba sorolhatók:

- tényekre, cselekedetre vonatkozóakra, amelyekre létezik objektív válasz, és az – elvileg – ellenőrizhető is;
- véleményekre, attitűdökre vonatkozó kérdések, amelyekre a válasz eleve szubjektív, időben, térben gyorsan változhat, és nem ellenőrizhető, legfeljebb bizonyos logikai összefüggések gondos tanulmányozása tárhat fel ellentmondásokat, az ellentmondás azonban nem tekinthető szükségszerűen rossz válasznak, szándékos torzításnak.

Mi kell ahhoz, hogy egy személy (család) a neki feltett különböző jellegű kérdésekre pontosan, jól válaszoljon? Ismét csak vázlatosan:

- akarjon válaszolni;
- megértse a feltett kérdést;
- tudja a választ, emlékezzék a lezajlott eseményekre (például vásárlás), cselekedetekre (például időmérés), illetve legyen véleménye a szóban forgó kérdéstről.

Ha a kikérdezés (az „interjú szituáció”) folyamatát belülről vizsgáljuk, azt látjuk, hogy a legjobb szándék mellett is sok olyan helyzet van, amikor a válasz „manipulálódik”, külső objektív és szubjektív tényezőktől függővé válik. Néhány tényező:

- a kérdések eltérő értelmezése;
- az emlékezet pontatlan;
- sokszor tudat alatti igyekezet a társadalmilag elvárt választ adni, amit természetesen a kérdező személye is óhatatlanul befolyásol;
- a válaszoló szinte mindig igyekszik önmagát kedvezően bemutatni;
- kényes kérdések kitérő, pontatlan, egyoldalú megközelítése stb.

Az említett problémák vázlatos átgondolása¹⁹ is jelzi, hogy egy lakossági adatfelvétel sikere nagymértékben a kérdőíven, a kérdések jó megfogalmazásán és a kérdező rátermettségén múlik. Mindezekből az is következik, hogy a felvétel végrehajtása után gondosan elemezni kell a tapasztalatokat, s értékelni azt: hogyan értelmezhetők az egyes kérdésekre adott válaszok.

¹⁸ A mintavételi hiba és a mintavételi terv összefüggéseinek módszertani kérdéseit vizsgálja például Kish, L. – Groves, R. M. – Krotki, K. P.: *Sampling errors for fertility surveys*. Occasional Papers. No. 17. ISI. WFS. Voorburg. Netherlands. 1976. 61 old.

¹⁹ Lásd bővebben: *Reprezentatív lakossági felvételek (összeírások, kikérdezések) végrehajtásának néhány módszertani kérdése*. KSH Utmutatók, segédletek, 11. Budapest, 1985. 41 old.; *Sudman, S. – Bradburn, N. M.: Response effects in surveys*. NORC. Aldine Publ. Co. Chicago. 1974. 257 old.

Ch. Cannell, az interjútechnika módszertanának neves szakértője szerint²⁰ a megfigyelési (survey) adatok valódisága attól függ, hogy sikerül-e a tudományosan kiválasztott személyeket rávenni arra, hogy részletes és pontos adatokat adjanak önmagukról, véleményükről és várakozásairól, anyagi körülményeik, jólétük megítéléséről, különböző tevékenységeikről, családi, társadalmi körülményeikről, iskolázottságukról stb. egy számukra idegen embernek, a kikérdezést végző összeírónak.

Az évek során az interjútechnika sokat fejlődött, de az alapvető cél változatlan maradt: meggyőzni a kiválasztott személyt, hogy közreműködjön, és megadja a lehető legteljesebb és legpontosabb információt. Ma már sokkal többet tudunk arról, miképpen érjük el ezt a célt, hogyan strukturáljuk a kérdőívet, a beszélgetést, hogy a lehető legjobb információkat kapjuk.

A válaszolók sokszor nem akarnak a számukra kényes kérdésben véleményüket nyilvánítani, de meg sem akarják tagadni a választ, ezért mondanak valamit, ami valószínűleg valamilyen irányba torzít. (Sokszor adódik a minden szép, jó, mi nem panaszkodunk szemlélet.)

Egyes szociológusok úgy érvelnek, hogy az ilyen jellegű válaszokat is elemezni, publikálni kell, mivel azok is bizonyos „realitást” jelentenek, nevezetesen azt, amit a lakosság erre a kérdésre a statisztikai hivatal összeírójának válaszol. Ez az érvelés és gyakorlat azonban nagy veszélyeket rejt magában, ami abból következik, hogy az összesített véleményekről egyáltalán nem lehet tudni, hogy azok valójában mit is jelentenek. Vannak ugyanis olyanok, akik minden bizonnyal komolyan, megfontoltan válaszoltak. Azokról viszont, akik mindegy, hogy milyen megfontolásból, de nem a tényleges véleményüket mondták meg, semmit sem tudunk. Nem tudjuk, hogy milyen mértékű az eltérés a valós helyzet és a kifejezésre jutott véleményük között.

A válaszadási, számbavételi hiba mérése meglehetősen nehéz feladat. Nem állnak rendelkezésre olyan egzakt matematikai statisztikai formulák, mint a mintavételi hiba becslésére. A szakirodalomban számos olyan vizsgálatról olvashatunk, amelyekben különböző kérdőívek, kérdezők, minta stb. azonos téma vizsgálatára vonatkozó eredményeit hasonlították össze, s abból vontak le következtetést a válaszadási hiba mértékére.

Az ELAR-felvételek körében egy ilyen irányú vizsgálatra került sor.²¹ 1983-ban az 1982. évi bruttó átlagkeresetet kérdeztük azonos személyektől két különböző időpontban, különböző felvétel keretében, a kérdést különbözőképpen megfogalmazva. Mint vártuk, a válaszolók viszonylag kis része, mindössze 33 százaléka sorolta magát nagy valószínűséggel a tényleges helyzetnek megfelelő jövedelem-kategóriába, míg 11 százaléka magasabb, 56 százaléka pedig alacsonyabb kategóriát jelölt meg. Amennyiben megengedünk egy kategóriányi eltérést, akkor már a válaszolók 78 százaléka adott többé kevésbé helyes választ. Ezek az eredmények azt jelzik, hogy óvatosan kell kezelni a jövedelemre vonatkozó adatokat azok gyakori, inkább lefelé történő torzítása miatt.

A reprezentatív felvételekből származó becslések megbízhatósága függ a megírók számától, mértékétől is. Módszertanilag eltérő megközelítést igényel, és következményeiben is mást jelent, ha egy tervezett kikérdezés például távollét vagy megtagadás miatt teljesen megghiúsul, vagy ha csak egy-egy kérdésre nem adtak választ.

²⁰ ISR Newsletter, The University of Michigan, 1985. Autumn, 4-5. old.

²¹ Lásd bővebben: Keleti András - Marton Ádám: Lakossági felvételek kereseti adatainak összehasonlító vizsgálata. Statisztikai Szemle, 1986. évi 11. sz. 1119-1125. old.

Az ELAR-felvételek körében a teljes meghiúsulások kezelését illetően kétféle gyakorlat alakult ki. A háztartásstatisztika esetében pótolni kell a kimaradó felvételt: a minták előírt terjedelmét biztosítani kell pótcímek igénybevételével. A többi felvételnél nincs pótlás: a minta terjedelme értelemszerűen csökken.

A meghiúsulások kezelésének legkézenfekvőbb módszere: mindent megtenni annak érdekében, hogy mértéke csak minimális legyen.

Viszonylag egyszerű a helyzet, ha távollét miatt hiúsul meg az összeírás. Ekkor többszöri későbbi felkereséssel pótolni lehet az összeírást. Amennyiben tartós távollétről, üres lakásról, nem létező címről stb. van szó, dönteni lehet arról, hogy újabb címek kijelölésével a hiányt pótoljuk-e, vagy sem.

Más a helyzet a megtagadások esetében. Az a megkérdezett döntésétől függ. Az összeíróval a kapcsolatfelvétel megtörtént, és a megkérdezett(ek) valamilyen okból úgy döntött(ek), hogy nem válaszol(nak) a felteendő kérdés(ek)re. Egy ilyen döntés hátterében sokféle szempont húzódnak meg. Bizalmatlanság, időhiány, rossz időpont, elzárkózás bizonyos magánjellegű információk feltárása elől stb. Mindezek azonban olyan tényezők, amelyek befolyásolhatók. Jó előkészítéssel, kellő propagandával, megfelelő felkéréssel, a felmérés társadalmi fontosságának megértésével, a kapott információk szigorúan bizalmas kezelésének hangsúlyozásával, a felkeresés időpontjának kellő megválasztásával stb. Az összeírók fellépésén, rátermettségén, rábeszélőkészségén nagyon sok múlik.

Vannak és mindig is lesznek olyanok, akiket semmilyen módon nem lehet válaszolásra bírni. Ahhoz, hogy a becslések megbízhatóságában a hiányzó adatok miatt bekövetkező csökkenést valamelyest értékelni lehessen, ismerni kellene a meghiúsulások megoszlását. Amennyiben ugyanis teljesen véletlenszerűen fordulnak elő, akkor lényegében csak arról van szó, hogy a minta terjedelme valamelyest csökkent. Ezt pedig a mintavételi hiba kiszámítása automatikusan figyelembe veszi.

Más a helyzet abban a szinte általánosnak tekinthető esetben, amikor a meghiúsulások, különösképpen a megtagadások megoszlása nem véletlen. Ennek felméréséhez azonban tudni kellene valamit a meghiúsult adatfelvételek mögött meghúzódó személyek társadalmi hovatartozásáról.

A meghiúsulások kedvezőtlen következményeinek csökkentésére számos módszer ismeretes, amelyeknek közös tulajdonsága, hogy valamilyen feltételezés segítségével pótolják (imputálják) a hiányzó adatokat.²² E módszerek ismertetése meghaladja e tanulmány kereteit.

A magyar gyakorlatban a felszorzási algoritmus automatikusan biztosítja, hogy a meghiúsult felvételek következtében „alábecslésre” ne kerüljön sor. Mint arról szó volt, a részleges meghiúsulások „pótlása” már a kérdőívek ellenőrzése során megtörténik, ami kétségtelenül egyszerű és „hatékony” eljárás, de egyáltalán nem nevezhető a legjobbnak. Szükséges lenne, hogy a jövőben vizsgáljuk és alkalmazzuk is a különböző imputálási technikákat, amelyekhez a feltételeket viszonylag egyszerűen meg lehetne teremteni.

A hiányzó adatok tekintetében nagyon találó a következő megállapítás: a legjobb megszerezni azokat közvetlenül az érdekelttől. A legrosszabb viszont nem tenni semmit.

TÁRGYSZÓ: Statisztikai mintavétel.

²² E módszereket a szakirodalom széles körben tárgyalja. Lásd például Kalton, G.: Compensating for Missing Survey Data. Report Series. SRC. ISR. The University of Michigan. 1983. 157 old.

РЕЗЮМЕ

Автор производит обзор программы, строения и функционирования действующей в венгерском Центральном статистическом управлении с 1983 года в модернизированном виде Единой системы сбора данных о населении. Постоянно возрастает потребность в информации о материальном, культурном и здравоохранительном состоянии населения, о характере происходящих внутренних процессов, ввиду чего в ходе истекших 5 лет имели место весьма загруженная программа и множество обследований. В 1988—1992 годы по различным соображениям следует ограничиться несколько более скромной программой.

Автор останавливается на факторах и причинах, которые воздействуют на функционирование Единой системы и в некотором смысле определяют достоверность полученных информации. Подчеркивает, что нужно уделять большое внимание планированию выборок, опросных листов и проведению обследований, следует тщательно анализировать возникновение срывов, проблемы в связи с толкованием отдельных вопросов (ошибочные ответы). В заключение излагает дальнейшие возможные направления методологических исследований и развития.

SUMMARY

The author reviews the programme, organizational structure and functioning of the Unified System of Household Surveys updated in 1983 in the Hungarian Central Statistical Office. More and more information should be collected on the financial, cultural, health etc. conditions of the population and on the characteristics of the inner processes of these spheres. This resulted in notable workload and significant number of surveys in the last 5 years. In the 1988—1992 period we must rest satisfied, due to various considerations, with a more modest programme.

The author deals with the factors, reasons which influence functioning of the System and determine, in certain aspect, the reliability of the information obtained. He stresses that much attention should be paid to the designing of samples, questionnaires and to conduct the survey. The failures and the interpretation problems of certain questions (reply error) should be analysed carefully. Finally, the possible future trends of research and development on methodology are outlined.