

KÖZPONTI STATISZTIKAI HIVATAL

# STATISZTIKAI MÓDSZERTANI FÜZETEK

AZ ELAR MINTA  
ÉS AZ 1984. ÉVI  
MIKROCENZUS MINTÁJÁNAK  
KIVÁLASZTÁSI ELJÁRÁSA

BUDAPEST, 1987

24

**STATISZTIKAI  
MÓDSZERTANI FÜZETEK  
24. szám**

**AZ ELAR MINTA  
ÉS AZ 1984. ÉVI  
MIKROCENZUS MINTÁJÁNAK  
KIVÁLASZTÁSI ELJÁRÁSA**

**Központi Statisztikai Hivatal  
Budapest, 1987**

1964. évi évkönyv

Készült:  
a Központi Statisztikai Hivatal  
Népesedésszatisztikai főosztályán

Főosztályvezető:  
Dr. Klinger András

Főosztályvezető helyettes:  
Dr. Kepecs József

1964. évi évkönyv

1964. évi évkönyv

1964. évi évkönyv

1964. évi évkönyv

Irtta:  
Éltető Ödön

ISSN 0231-0554  
ISBN 963 340 623 4

# TARTALOM

	Oldal
ÖSSZEFOGLALÓ -----	5
I. AZ ÚJ ELAR MINTA	
1. Az új általános célu mintakeret kialakításának háttere, alapelvei	
1.1 Miért volt szükség új ELAR mintára? -----	7
1.2 Elvárások az új mintával szemben -----	8
1.3 Az új ELAR minta kiválasztásával kapcsolatos néhány alapelv -----	9
1.4 Az új ELAR minta integráltsági foka -----	12
1.5 A minta területi elosztása -----	13
2. Az arányos minta felépítése	
2.1 Mintanagyság -----	15
2.2 Rétegzés -----	16
2.3 A települések kiválasztása és a település-minta ellenőrzése -----	20
2.4 A körzetek kiválasztása és a körzet-minta ellenőrzése -----	24
2.5 A minta kiegészítése új lakótelepi körzetekkel -----	28
2.6 A minta arányosságának és a lakótelepi reprezentációnak ellenőrzése -----	28
3. A működő nem arányos ELAR minta kiválasztása	
3.1 A nem arányos minta kialakítása -----	30
3.2 Az arányos és nem arányos minta összetétele -----	32
3.3 Két rész-minta kialakítása -----	33
3.4 Sulyozási algoritmus -----	34
3.5 Az új ELAR minta ellenőrzése -----	36
3.6 A minta karbantartása, rotáció -----	40
II. AZ 1984. ÉVI MIKROCENZUS MINTÁJA	
1. A mintavétel alapelvei -----	43
2. Települések és számlálókörzetek kiválasztása -----	44
3. A mikrocenzus mintájának összetétele -----	46
4. Lakások kiválasztása -----	49
5. Intézeti háztartások kiválasztása -----	49
6. A minta ellenőrzése és korrekciója -----	50
PE3IOME -----	53
SUMMARY -----	55

## ÖSSZEFOGLALÓ

---

A füzet elsősorban az Egységes Lakossági Felvételi Rendszer /ELAR/ keretében 1983. január 1-től 1992 végéig végrehajtásra kerülő reprezentatív felvételek mintáihoz mintavételi keretül szolgáló, népszámlálási körzetekből álló új általános célú minta kiválasztási módszerét ismerteti. Mivel azonban az 1984 októberében végrehajtott 2 %-os mikrocenzus mintája szoros kapcsolatban van az új ELAR mintával, a füzet ennek mintavételi módszere ismertetésére is kitér.

1976 és 1982 között végrehajtott ELAR felvételek két, még 1972-ben kiválasztott független rész-mintán alapultak. Az első fejezet azt tárgyalja, miért volt szükség egy új, egységes általános célú mintára, mik voltak az elvárások az új mintával szemben, milyen főbb hasonlóságok és eltérések vannak a régi és az új ELAR minta kiválasztási módszere között. Az eltérések között a legfontosabb, hogy a régi mintával ellentétben az új ELAR minta területileg nem arányos elosztású: Budapest és a nagyobb városok, illetve az inaktív jellegű települések, körzetek tényleges súlyuknál kisebb, a többi település, különösen a paraszti jellegű települések pedig nagyobb arányban szerepelnek a mintában.

A működő nem arányos minta kiválasztása két fázisban történt: elsőnek egy, az 1980. évi népszámlálási körzetek 15 %-át tartalmazó arányos minta került kiválasztásra. A füzet 2. fejezete ennek felépítését, kiválasztási módszerét ismerteti: a mintanagyság meghatározását, a települések, illetve körzetek rétegzésének módját, kiválasztási eljárását, a település-, illetve körzetminta ellenőrzését. Az 1980. évi körzetek 15 %-os mintáját még ki kellett egészíteni az 1980-1982 években épült új lakótelepi körzetek megfelelő arányú mintájával.

A működő ELAR minta kiválasztása ebből a 15 %-os arányos mintából történt települések, illetve körzetek nem azonos arányban történő elhagyásával. A 3. fejezet ennek módját tárgyalja bemutatva egyúttal az arányos és nem arányos minta összetételét. Az új ELAR minta is tartalmaz két almintát, de csak a körzetek szintjén, a tele-

püléseket tekintve egyetlen integrált minta van. A 3. fejezet ismerteti még a két alminta kialakításának módszerét, a felvételek feldolgozásánál alkalmazandó sulyozási algoritmust, valamint ennek és a minta reprezentativitásának ellenőrzésére szolgáló, még a minta gyakorlati alkalmazása előtt végzett teszt számítások eredményeit. A fejezet utolsó pontja a minta karbantartásával és a rotációval kapcsolatos kérdéseket tárgyalja.

A füzet II. része az 1984. évi mikrocenzus mintájának kiválasztási módszerét ismerteti. A mikrocenzus egy arányos elosztású 2 %-os lakásmintán alapult, kiegészülve az intézeti háztartások ugyanolyan arányú mintájával. A magánháztartásokra vonatkozó minta alapját az ELAR minta első fázisában kiválasztott 15 %-os arányos körzetminta alkotta. Annak érdekében azonban, hogy a mikrocenzust minél nagyobb arányban a jól képzett és nagy gyakorlatra szert tett ELAR kérdőbiztosok hajtsák végre, a 15 %-os mintát a 15 ezernél kisebb települések tekintetében tömörítettük. A füzet részletesen ismerteti a tömörítés módját a kisebb és nagyobb településeken, a mikrocenzus mintájának összetételét, a lakások, illetve intézeti háztartások kiválasztásánál követett eljárást, valamint a minta ellenőrzését.

# I. AZ ÚJ ELAR MINTA

## 1. AZ ÚJ ÁLTALÁNOS CÉLU MINTAKERET KIALAKÍTÁSÁNAK HÁTTERE, ALAPELVEI

### 1.1 MIÉRT VOLT SZÜKSÉG ÚJ ELAR MINTÁRA?

1972-ben került sor először Magyarországon ún. általános célu minta kialakítására. Ez a mintakeret az 1970. évi népszámlálási számlálókörzetek 10 %-át ölelte fel, kiegészülve az 1970-1972 években épült új, lakótelepi körzetekből vett 10 %-os mintával. Az általános célu minta kiválasztása azzal a céllal történt, hogy az elkövetkező 10 évben, 1973-1982-ben mintavételi keretül szolgáljon az összes, a lakosság körében végrehajtásra kerülő reprezentatív felmérésekhez. Így mindjárt az 1973. évi mikrocenzus, illetve a jövedelmi és átrétegződési felvételek mintái is ebből a 10 %-os mintakeretből kerültek kiválasztásra. Az állandó mintakeret előnyei azonban igazából csak azután érvényesültek, hogy 1976. január 1-től működni kezdett az állandó kérdőbiztosi gárdával és munkájukat irányító megyei szervezetekkel az Egységes Lakossági Adatfelvételi Rendszer. Az ELAR abban az időben két - a minta és a közreműködő kérdőbiztosok tekintetében - egymástól független részből állt /ELAR I. és ELAR II./. A két minta az 1972-ben kiválasztott általános célu minta egy-egy 20 %-os, azaz a népszámlálási körzetek 2-2 %-át felölelő részmintája volt. Ezek a részminták már az elsődleges mintavételi egységek, azaz a települések /illetve nagyobb városokban a körzetek/ szintjén váltak egymástól.

A népszámlálási körzetek egy rögzített mintáján alapuló általános célu minta számos előnye mellett egy idő után elkerülhetetlenül elavul, illetve kimerül, egyre kevésbé képes megfelelni eredeti céljának. Ez még fokozottabban áll az egyes részmintákra. Ennek alapvetően két oka van: egyrészt a rögzített minta nem volt képes elég flexibilisen követni az ország lakásállományában, illetve a lakosság belső vándorlásában és átrétegződésében tíz év alatt végbement jelentős változásokat, illetve ennek csak egy részét, a lakótelepi lakások számának növekedését lehetett a mintának évenként megfelelő számú új lakótelepi körzettel történő kiegészítése révén

nyomonkövetni, másrészt az ELAR működésének első hét évében jelentős számban végrehajtott különböző megfigyelésekbe bevont háztartások számos mintakörzet esetén lassanként kimerítették a körzet teljes lakásállományát. Ez utóbbi különösen az ELAR II. mintájánál jelentett problémát, a háztartásstatisztikai megfigyelésnél ui. a különböző lakossági rétegek nem arányosan kerültek kiválasztásra, a közreműködést nem vállaló háztartások helyett pedig azonos nagyságu és rétegű új háztartásokat jelölnek ki a mintába. Mindez indokoltá tette, hogy - az eredeti elképzeléseknek megfelelően - az 1980. évi népszámlálás után, annak anyagára támaszkodva új általános célu minta kiválasztására kerüljön sor.

Más oldalról szervezési szempontok is indokoltá tették egy új minta kiválasztását. Az ELAR működése során néhány év alatt összegyűlt tapasztalatok arra engedtek következtetni, hogy két, egymástól jórészt már a települések szintjén elváló rész minta és ennek megfelelően két kérdőbiztosi gárda fenntartása nem teszi lehetővé az ELAR szervezet megfelelő hatékonyságu működtetését. Egy új, integrált szervezet viszont csak új, egységes minta alapján jöhetett létre.

## 1.2 ELVÁRÁSOK AZ ÚJ MINTÁVAL SZEMBEN

A fentiek alapján a régi két független rész minta helyett egyetlen, egységes ELAR mintára volt szükség. Ennek az új "általános célu" mintának kell folyamatos karbantartás mellett 10 éven keresztül - 1983-tól kezdve 1992 végéig - keretül szolgálnia az ELAR keretében végrehajtásra kerülő különböző ismétlődő, illetve egyszeri háztartási és egyéb megfigyelésekhez, így elsősorban a kétévenként rendszeresen sorra kerülő társadalmi-demográfiai alapadat-felvételhez /korábban munkaerő felvétel/ és a háztartási költségvetési felvételhez, az ötévenkénti jövedelmi felvételekhez, valamint más demográfiai, társadalomstatisztikai, egészségügyi, stb. felmérésekhez. Az új ELAR mintának, mint mintavételi keretnek olyannak kellett lennie, hogy a belőle kiválasztott minták - a minta nagyságától függő pontossággal - reprezentálják az ország - esetenként a nagyobb területi egységek, régiók, megyék - lakásait, háztartásait, illetve népességét vagy egyes népességcsoportjait. Követelmény volt továbbá, hogy a mintavételi keretül szolgáló ELAR minta elegendő számú lakáscimet, illetve

háztartást tartalmazzon a címek /háztartások/ megfelelő mértékű többszöri rotálásához /részleges vagy teljes cseréjéhez/ a szóbanforgó tíz év alatt, illetve a háztartási költségvetési felvételnél még a kieső háztartások ugyanazon társadalmi rétegből történő pótlásához is.

Az 1980. évi népszámlálás anyagából kiválasztott mintát már az induláshoz ki kellett egészíteni az 1980-82 években újonnan épült lakások megfelelő nagyságu mintájával és ezt a kiegészítést, illetve cserét évenként el kell végezni azért, hogy az aktuális minták mindig a tényleges lakásállományt reprezentálják.

### 1.3 AZ ÚJ ELAR MINTA KIVÁLASZTÁSÁVAL KAPCSOLATOS NÉHÁNY ALAPELV

a/ Az új minta kiválasztásánál is - miként 1972-ben az első általános célu minta esetében is - gyakorlatilag csak többlepcsős kiválasztási eljárás jöhetett szóba. Egy lépcsőzés nélkül kiválasztott minta - azonos mintanagyság mellett - általában megbízhatóbb eredményeket biztosít, s elvileg lett is volna lehetőség a népszámlálás anyagából közvetlenül lakásokat, illetve háztartásokat kiválasztani. Ilyen mintavétel esetén azonban gyakorlatilag az ország majd minden településéről kerülnének be lakások a mintába, így az összeíróknak meglehetősen sokat kellene utazniuk, ami rendkívül megnövelné a felvételek költségeit. Ma még - s feltehetően a közeljövőben is - a kérdőbiztosok zöme csak saját lakóhelyén - esetleg lakóhelyéhez közeleső és könnyen megközelíthető településen - vállalja az összeírói munkát. Így a mintának szükségszerűen bizonyos számú településre kellett koncentrálnia, azaz a nagyobb, ún. önreprezentáló városoktól eltekintve a településeket kellett elsődleges mintavételi egységnek tekinteni.

Másodlagos mintavételi egységekként körzetek kiválasztását, azaz a mintának a kiválasztott településeken belüli további koncentrációját ugyancsak gyakorlati szempontok tették szükségessé. A felvételekkel járó munka - a kiválasztott címek felkeresése - is több lenne ugyan, ha településeken belül már közvetlenül lakásokat választanánk ki, de különösen a minta folyamatos karbantartása - a teljes címanyag évenkénti bejárása, ellenőrzése - jelentene rendkívül nagy többlet terhet a kérdőbiztosok számára. Ezért feltétlenül szükséges-

nek látszott, hogy a kiválasztott települések zömében - a legkisebb települések kivételével - további mintavételi egységekként népszámlálási számlálókörzetek kerüljenek kiválasztásra. Egyébként a számítások azt mutatták, hogy a körzetek megfelelő rétegzésével és az egy-egy településről kiválasztandó körzetek számának emelésével a körzetek kiválasztása nem csökkenti számottevően az adatok megbízhatóságát.

Gyakorlati megfontolások és az ELAR szervezet és felvételek működéséről 1976 óta szerzett tapasztalatok alapján tehát célszerűnek mutatkozott az új ELAR mintát is a települések zömét tekintve három lépcsőben - település, körzet, lakás - kiválasztani. A három lépcső az egyes konkrét mintákra vonatkozik, maga a mintavételi keretet alkotó általános célú minta ennek megfelelően két lépcsőben került kiválasztásra.

b/ 1972-ben az első általános célú minta kiválasztásakor részletesen megvizsgáltuk a különböző lehetséges többlépcsős mintavételi eljárásokat, azok következményeit, előnyeit, hátrányait.<sup>1/</sup> A különböző elméleti és gyakorlati szempontokat figyelembevéve az látszott legelőnyösebbnek, ha mind a települések, mind azokon belül a körzetek - a lakások számával mért - nagysággal arányos valószínűséggel kerülnek kiválasztásra, a kiválasztott körzeteken belül pedig az egyes minták azonos számú lakást tartalmaznak. Mind az 1973. évi mikrocenzus és a hozzá kapcsolódó felvételek, mind az ELAR keretében 1976 óta végrehajtott folyamatos és egyszeri háztartási felmérések tapasztalatai messzemenően alátámasztották az akkori döntés helyességét, különösen a területi munka szervezését teszi lényegesen egyszerűbbé az összeírók egyforma terhelése. Ezért minden a mellett szólt, hogy az új ELAR minta kiválasztásánál is ezt a mintavételi eljárást célszerű alkalmazni.

c/ Alapelvnek tekintettük, hogy a minta reprezentativitásának növelése érdekében mind a településeket, mind azokon belül a körzeteket rétegzett mintavételi eljárással célszerű kiválasztani. Az alapelv természetesen még nem ad választ arra a kérdésre, miként, milyen ismérvek vagy ismérvkombinációk alapján célszerű rétegezni, ezzel a későbbiekben még majd részletesebben foglalkozunk.

1/ Részletesen l. "Az általános célú háztartási minta kialakítása" Statisztikai Szemle, 50. évf. 10. szám /1972 október/.

d/ Az 1972-ben kiválasztott általános célu minta ugy épült fel, hogy a mintába került 5 000-nél kisebb lélekszámú községek mind-egyikéből 2 számlálókörzetet tartalmazott a minta, a nagyobb településeknél pedig a számlálókörzetek 10 %-át. A gyakorlat azonban bebizonyította, hogy az ismétlődő háztartási megfigyelések számára a rotációt is figyelembe véve még 10 évnél rövidebb időre sem nyujt megfelelő mintavételi keretet a községenkénti 2 körzet, a cimanyag elég gyorsan kimerül. Emellett 2 körzet általában nem tudja elég jól reprezentálni a községek lakosságát, sem a lakásállományban, illetve társadalmi rétegződésben bekövetkező jelentős változásokat. Célszerűnek tűnt ezért az új ELAR mintát úgy kiválasztani, hogy a mintába kerülő községekben bővebb mintavételi keret - több lakáscim - álljon rendelkezésre a különböző háztartási megfigyelésekhez, ami településenként több körzet kiválasztását tette szükségessé.

A mintakörzetek nagyobb száma természetesen nem jelenti azt, hogy az egyes háztartási megfigyelések mintáit is növelni kell, a konkrét minták nagysága ettől függetlenül határozható meg. A kérdőbiztosoknak lényegében a minta évenkénti karbantartása - a körzetek teljes cimbejárása /részletesebben 1. a 3.5 pontban/ - szempontjából jelent a több körzet valamelyest munkatöbbletet. Ugyanakkor a körzetek megfelelő rétegzésével kiválasztott több körzet jobb lehetőséget nyujt a község lakosságának reprezentálására, s így ez a változtatás - amellett, hogy gyakorlati szempontok miatt szükséges is volt - hozzájárult az egyes minták eredményeinek megbízhatóságához.

e/ Nem merült fel indok annak a gyakorlatnak a megváltoztatására, hogy a végső mintavételi egység a lakás legyen, függetlenül attól, hogy a megfigyelés lakásokra, háztartásokra vagy az azokban élő bizonyos személyekre vonatkozik. Ettől a gyakorlattól csak a családi költségvetésekre vonatkozó felmérés tér el, ahol meggyűlés esetén azonos társadalmi rétegből történő pótlás megvalósíthatósága érdekében a mintakörzetekben lakó háztartásokat egy előzetes felmérés adatai alapján előbb társadalmi rétegekbe sorolják és a kiválasztás a már rétegenként csoportosított háztartások listájából külön-külön történik.

A fenti alapelvek természetesen még nem határozták meg egyértelműen az új ELAR minta kiválasztásánál követendő mintavételi eljárást, a minta felépítését. A kiválasztás gyakorlati megvalósításához még dönteni kellett olyan fontos kérdésekben, mint amilyen a

minta integráltsági foka, illetve a minta arányos vagy nem arányos elosztása.

#### 1.4 AZ ÚJ ELAR MINTA INTEGRÁLTSÁGI FOKA

Az Egységes Lakossági Adatfelvételi Rendszer 1976 és 1982 között két részmintán alapult, amelyek közül az egyik minta a folyamatos munkaerő felvétel és a kapcsolódó társadalomstatistikai és egyéb felvételek céljaira szolgált, a másik pedig a szintén folyamatos háztartási költségvetési felvétel céljára. A két rész minta az 1972-ben kiválasztott általános célu minta egy-egy 20 %-os, azaz a népszámlálási számlálókörzetek 2-2 %-át felölelő almintája volt, amelyek már az elsődleges mintavételi egységek szintjén függetlenek voltak, tehát a nagyobb városoktól eltekintve különböző településeket foglaltak magukba. Ennek következtében a kérdőbiztos hálózat is különböző volt. Az egységességet egyrészt a két rész minta teljesen azonos strukturája, kiválasztási módja jelentette, másrészt az egységes irányító szervezet a megyei KSH Igazgatóságokon. E mellett voltak olyan háztartási megfigyelések - mint pl. az 1978 elején végrehajtott jövedelmi felvétel -, amelyek mindkét részmintára kiterjedtek, amelyeknél tehát az ELAR mint integrált rendszer működött.

A tapasztalatok azt mutatták, hogy az egész rendszer hatékonyabb működtetéséhez az integráció magasabb foka szükséges, ami első sorban egységes, egyetlen kérdőbiztos hálózat kialakításának igényeként jelentkezett. Ez természetesen magával vonta, hogy a települések szintjén egyetlen mintára volt szükség. Ilyen integráció nemcsak a kérdőbiztos hálózat jobb, gazdaságosabb felhasználását teszi lehetővé, hanem egyúttal a mintákból nyert adatok nagyobb megbízhatóságát is eredményezi, mivel így a minta több elsődleges mintavételi egységet tartalmaz. Elvileg az integráció olyan foka is elképzelhető, hogy minden felvételt egyetlen mintán hajtanak végre, s csak az évenkénti rotáció során változik a minta összetétele. Ez azonban a felvételekbe bevont háztartások túlzott megterhelését vonná maga után. Különösen a háztartási költségvetési felvétel jelent meglehetősen sok munkát a közreműködő háztartások számára, ezért a közreműködés megtagadása arányának nem kívánatos növekedését elkerülendő mindenképpen célszerű, ha az ebben a felvételben közreműködő háztartásokat nem terheljük további felvételekkel. Mindez azt jelenti, hogy az új ELAR mintának,

mint mintakeretnek olyannak kellett lennie, hogy egyszerre legalább két háztartási /lakás/ minta kiválasztását tegye lehetővé.

Mint említettük, a háztartási költségvetési felvételnél a körzeten belüli kiválasztás - a többi ELAR felvételtől eltérően - nem a körzet lakásainak címjegyzékéből történik, hanem a háztartások társadalmi rétegek szerint csoportosított jegyzékéből. Ez viszont azt jelenti, hogy több szempontból is bonyodalmakat jelentene, ha a háztartási költségvetési felvételre szolgáló mintát és a többi felvételhez szükséges lakásmintát ugyanazokból a körzetekből kellene kiválasztani. Így az látszott legcélszerűbb megoldásnak, ha a települések szintjén egységes új ELAR minta két azonos számú, de különböző körzeteket tartalmazó részmintából tevődik össze. Ez a megoldás biztosítja az integráció legfőbb előnyeit - a településeket tekintve egyetlen, s ezért több települést felölelő minta, egységes kérdőbiztosítási hálózat -, a teljes integrációval járó hátrányok és nehézségek nélkül.

#### 1.5 A MINTA TERÜLETI ELOSZTÁSA

Mind az ELAR létrehozása előtt végrehajtott reprezentatív háztartási felvételek zömének mintája, mind az 1976-tól 1982-ig működő első ELAR minta arányos minták voltak abban az értelemben, hogy a kiválasztási arány a különböző rétegekben, területi egységeknél egyforma volt. Ettől a gyakorlattól csak az 1973. évi jövedelmi felvétel tért el, ahol Budapesten és a négy legnagyobb megyében a kiválasztási arány fele volt a többi megyében alkalmazottnál. A háztartási költségvetési felvételnél ugyan társadalmi rétegek szerint különböző volt a kiválasztási arány, de ez a körzeteken belüli eltérő arányú kiválasztás révén valósult meg és társadalmi rétegeken belül területileg e felvétel mintája is arányos elosztású volt.

Az új ELAR minta kiválasztásánál alkalmazandó mintavételi eljárás meghatározásánál is felmerült a kérdés és dönteni kellett arról, hogy arányos vagy nem arányos elosztású legyen-e az új minta. Az arányos minta előnye az egyszerűbb kiválasztás és adatfeldolgozás. Arányos minta esetén ugyanakkor a minta szükségtelenül nagy része jut Budapestre, illetve a nagy városokra, aminek következtében kevesebb településre terjedhet ki a minta, s így az adatok megbízhatósá-

ga csökken, mert a mintavételi hiba jelentős része éppen a települések közötti szóródásból ered. Vannak továbbá olyan háztartási felvételek - elsősorban a háztartási költségvetési felvétel - amelyeknél követelmény, hogy főbb társadalmi rétegenként is közöljön adatokat a hivatal, s mivel egyes rétegek súlya elég kicsi, arányos minta esetén e rétegek súlya a mintában is hasonlóan kicsi lenne, azaz csak rendkívül nagy mintával lenne elérhető, hogy e rétegekbe is annyi megfigyelés jusson, ami biztosíthatja az adatok közölhetőségét. Társadalmi rétegenként nem arányos minta viszont mérsékelt mintanagyság mellett is eleget tud tenni e követelményeknek.

Még egy további szempont szölt a nem arányos minta mellett. Egy adott időpontra arányos elosztású minta bizonyos idő múlva már nem lesz arányos, hiszen a különböző típusú településeken nem azonos arányú a lakásépítések és megszünések száma, így ahhoz, hogy a mindenkori minta arányos maradjon, időről-időre korrigálni kell oly módon, hogy bizonyos településeken megszünik vagy csökken a minta, más helyeken nő, illetve esetleg újabb településeket kell bevonni a mintába. Ez újra meg újra szervezési problémákat jelentene, különösen akkor, ha pénzügyi korlátok miatt a minta méretét nem lehet folyamatosan növelni. Nem arányos minta esetén ugyanakkor nem jelentkezik ez a probléma, hiszen csak a súlyszámokat kell folyamatosan módosítani a lakásállomány strukturájának változása szerint.

A különböző szempontokat és igényeket mérlegelve végül olyan döntés született, hogy az új ELAR minta nem arányos elosztású legyen. A minta kialakításánál az arányos elosztástól két szempont szerint látszott célszerűnek eltérni:

- annak érdekében, hogy a minta - adott mintanagyság mellett - minél több települést öleljen fel, Budapesten és a nagyobb városokban az átlagosnál kisebb volt a kiválasztási arány, a községekben pedig nagyobb;
- a községek és körzetek alkalmas rétegzésével és az így kialakított rétegekre vonatkozó különböző arányú kiválasztással igyekeztünk elérni egyes kisebb létszámú, de fontos lakossági rétegek - mint pl. a termelészövetkezeti parasztság - jobb reprezentációját a mintában.

Mivel mind a nagyvárosi népesség aránya, mind a társadalmi rétegek strukturája eltérő az egyes megyékben, a fenti két szempont szerinti nem arányos kiválasztás egyúttal maga után vonta, hogy az egyes megyékre vonatkozóan is kisebb-nagyobb mértékben különböző volt a kiválasztási arány.

Részben a kiválasztás könnyebb technikai lebonyolítása érdekében, részben pedig az arányos mintát igénylő, 1984-re tervezett mikrocenzus mintájának előkészítéseképpen célszerűnek látszott az új ELAR mintát két fázisban kiválasztani: először egy arányos elosztású számlálókörzeti mintát, majd ebből a célul kitűzött nem arányos mintát.

## 2. AZ ARÁNYOS MINTA FELÉPÍTÉSE

### 2.1 MINTANAGYSÁG

Általános célú mintakeret kiválasztásakor a mintanagyság kérdése másképp merül fel, mint egy konkrét felvételhez kiválasztandó minta esetén. A mintakeret nagysága - a minta által felölelt számlálókörzetek száma - közvetlenül nem határozza meg az egyes felvételek mintanagyságát, ez utóbbi a felvétel céljától, az adatok részletességére, megbízhatóságára vonatkozó igénytől, stb. függően adott nagyságú mintakeret mellett is még elég tág határok között változhat. Az arányos minta nagyságával kapcsolatban két dolgról kellett dönteni: a/ a népszámlálási körzetek hány százalékát ölelje fel a mintakeret; b/ az adott számú körzetet hány településről válasszuk ki.

Az 1972-ben kiválasztott első általános célú minta a körzetek 10 %-át ölelte fel oly módon, hogy az 5 000-nél kisebb lélekszámú településekről 2 számlálókörzet került a mintába. Mint már említettük, az első ELAR minta tapasztalatai azt mutatták, hogy a tervezett tíz évre két körzet cimanyaga a szükséges rotációt is figyelembe véve nem minden esetben biztosítja a rendszer zavartalan működését. Az alapelvek tisztázása során az is egyértelművé vált, hogy az új minta a településeket tekintve egységes kell legyen és a körzetek szintjén két egyforma és egyenlő nagyságú részmintából kell álljon.

Mindez azt jelentette, hogy a települések zömében részmintánként legalább 3, összesen 6 körzetet kell tartalmaznia az új mintakeretnek. Ez már maga után vonta, hogy - hacsak nem akartuk a mintát a megbízhatóság rovására viszonylag kevés számú településre korlátozni - az arányos mintának körzetek több mint 10 %-át kell felölelnie. Másrészt a számítások azt mutatták, hogy mind a munkaerő felvétel és a kapcsolódó egyéb felvételek esetében, mind a háztartási költségvetési felvétel céljára a körzetek kb. 4 %-ának cimanyagára van szükség ahhoz, hogy a tízéves időszak alatt a szükséges rotációkhoz, illetve pótlásokhoz elegendő cím álljon rendelkezésre. Mivel a nem arányos kiválasztás az előzetes tervek szerint kb. felére redukálja az arányos mintát, végeredményben az adódott, hogy az arányos mintának a népszámlálási körzetek 15 %-át kell tartalmaznia ahhoz, hogy a végső minta a fenti követelményeknek eleget tudjon tenni.

A kiválasztandó települések számát egyrészt az szabta meg, hogy célszerűnek mutatkozott, hogy az arányos minta tartalmazza az összes 10 ezernél népesebb települést, másrészt a kisebb települések esetén az arányosság, a 15 %-os körzetminta és a településenként kiválasztandó 6 körzet már meghatározta a mintába kiválasztandó települések számát. A kis községek esetén, amelyek hatnál kevesebb körzetet tartalmaztak, több - két vagy három - ilyen kis települést tekintetünk egy mintavételi egységnek, azaz a mintába kerülő települések számát nem egyszerűen a körzetek 15 %-ának hatodrésze adta, ez csak az 1 500 és 10 000 közötti lélekszámú településekre volt érvényes.

Végeredményben így az arányos minta kiterjedt a 137 /az 1980. évi népszámlálás adatai alapján/ 10 ezernél népesebb településre, valamint 561 ennél kisebb településre, azaz összesen 698 településre, ezek közül 591 volt község, a többi város.

## 2.2 RÉTEGZÉS

A szükségszerűnek bizonyult többlépcsős kiválasztási eljárásból eredő nagyobb mintavételi hibát a települések és a körzetek megfelelő rétegzésével igyekeztünk ellensúlyozni. A rétegzés ugyanis általában csökkenti a mintavételből eredő hibát, azaz növeli a mintából kapott eredmények megbízhatóságát.

1972-ben az általános célu minta kiválasztásakor az 5 000-nél kisebb lélekszámú községeket hat nagyságkategóriába, valamint - a mezőgazdasági népesség aránya alapján - három foglalkozási rétegbe soroltuk. Ezenkívül, mivel a kiválasztás során az egyes rétegekbe tartozó községeket megyék, azon belül járások szerint rendeztük sorba, ez automatikusan egy területi rétegzést is biztosított. E községeken belül a körzetek kiválasztásánál nem történt rétegzés, az 5 000-nél népesebb településeken - ezek mind szerepeltek az általános célu mintában - viszont a körzeteket kiválasztás előtt 9, a foglalkozási struktúra szerint képzett rétegbe soroltuk.

A tapasztalatok alapján, illetve az időközben bekövetkezett változások miatt célszerűnek mutatkozott az új ELAR minta kiválasztásánál bizonyos mértékben módosítani a rétegek meghatározását, illetve új szempontokat is figyelembe venni a rétegek kialakításánál:

- az 1972. évi általános célu mintától eltérően a 15 %-os arányos minta kiválasztásánál csak a 10 ezernél népesebb településeket tekintettük önreprezentálóknak, azaz önálló rétegeknek, amelyek mindegyikéből történt körzet kiválasztás. /Egyébként az 1972-es minta almintái, s így az 1976-tól működő ELAR I. és II. minták már nem tartalmazták az összes 5 000-nél népesebb települést/;
- az új ELAR minta kiválasztásánál már nem volt indokolt a községeket csupán a mezőgazdasági népesség aránya alapján rétegezni. E helyett a társadalmi rétegek szerinti struktúra valamilyen komplexebb figyelembe vétele látszott célszerűbbnek;
- tekintettel arra, hogy a mintavételi terv egy-egy mintába kerülő 1 500 - 10 000 lakosu településből az 1972-es 2 helyett 6 körzet kiválasztását írta elő, lehetővé vált és célszerűnek látszott minden településen a körzeteket is rétegezni, azért, hogy a kiválasztandó 6 körzet lehetőleg jól reprezentálja az egyes mintába került települések lakosságát;
- az inaktív népesség 10 év alatt bekövetkezett nagyarányú növekedése indokoltta tette, hogy a körzetek rétegeinek kialakításánál ne csak az aktív népesség foglalkozási strukturáját, hanem az inaktív népesség arányát is figyelembe vegyük;

- továbbra is lényegesnek tűnt a külterületi körzetek külön rétegeként való kezelése.

A fenti szempontok természetesen még nem határozták meg egyértelműen sem a rétegek számát, sem azok tartalmát. Ez vonatkozott különösen a települések, illetve körzetek foglalkozási struktúrára, illetve társadalmi réteg szerinti rétegeinek kialakítására. Nagyság szerinti besoroláshoz kisebb módosítással az 1972-es minta kiválasztásánál alkalmazott kategóriákat tartottuk meg. Így az alkalmazott nagyságkategóriák az alábbiak:

Település lélekszáma	Réteg jelzőszáma	Település lélekszáma	Réteg jelzőszáma
- 499	1	5 000- 9 999	6
500- 999	2	10 000-14 999	7
1 000-1 499	3	15 000-24 999	8
1 500-2 999	4	25 000-	9
3 000-4 999	5		

A kategóriák a településeknek az 1980. évi népszámláláskor regisztrált lakónépessége alapján kerültek kialakításra. A 7-9 nagyságkategóriák megkülönböztetésére természetesen az arányos minta kiválasztásához nem volt szükség, ezeknek csak a mintavétel második fázisában, a nem arányos minta kialakításakor volt szerepük.

A települések, illetve körzetek jelleg, valamint társadalmi-foglalkozási összetétel szerinti rétegeinek kialakításánál elsősorban azt kellett szem előtt tartani, hogy a mintavétel második fázisában a nem arányos kiválasztás egyik szempontja a kis súlyú társadalmi rétegek jobb reprezentációja volt. Konkréten az volt a cél, hogy a nem arányos kiválasztás révén az alapvető osztályok és rétegek közül a szövetkezeti parasztság és a kettősjvedelmű réteg számarányánál nagyobb, az inaktív háztartások pedig kisebb súllyal szerepeljenek a végső, nem arányos mintában. Kézenfekvő lett volna ennek alapján a körzeteket, illetve településeket a bennük élő háztartások társadalmi rétegek szerinti megoszlása, illetve a szövetkezeti paraszti, kettősjvedelmű és inaktív háztartások aránya alapján rétegezni. Ezt azonban két ok miatt nem lehetett megvalósítani: az egyik inkább technikai jellegű volt, a másik inkább tartalmi. A mintavétel előkészítéséhez, így speciálisan a rétegek kialakításához az 1980.

évi népszámlálás adatszalai álltak rendelkezésre. Ezek nem tartalmazták a háztartások alapvető osztályokba és rétegekbe való besorolását, csak a keresőkét. Elvileg ugyan meg lehetett volna oldani az adatszalg kiegészítését a háztartások besorolásával, gyakorlatilag azonban ez olyan időeltelődást eredményezett volna a mintavétel előkészítésénél, hogy el kellett tekinteni tőle. Tartalmilag a nehézséget az jelentette, hogy 1980-ban már sem a körzetek, de még kevésbé a települések között gyakorlatilag nem léteztek tiszta paraszti vagy tiszta inaktív körzetek vagy települések, minden településen, de szinte minden körzetben is vegyesen éltek különböző társadalmi rétegekhez tartozó háztartások. Így csak az a megoldás maradt, hogy a keresők osztály- és rétegbesorolása alapján próbáljuk közelíteni a kívánt célt, olyan rétegeket kialakítani, amelyek egyrészt sok szempontból homogénebbek, mint a teljes sokaság, s így e rétegzés elősegíti, hogy a mintából megbízható becsléseket kapjunk, másrészt alapul szolgálhatnak a mintavétel második fázisában a megvalósítani szándékozott nem arányos kiválasztáshoz.

Mind a komplexebb rétegzés, mind a későbbi nem arányos kiválasztás indokolta, hogy az 1972-es általános célu minta kiválasztásánál követett eljárástól eltérően a 10 ezernél kisebb lélekszámú településeket és a belterületi körzeteket azonos kritériumok alapján soroljuk jelleg szerinti rétegekbe. Végül a belterületi körzeteket, illetve a 10 ezernél nem népesebb településeket az alábbi jelleg szerinti rétegekbe soroltuk:

1. réteg: a munkásosztályhoz tartozó aktív keresők abszolút többségben vannak az inaktív és inaktív keresők között /tiszta munkás jellegű körzetek, illetve települések/
2. réteg: az aktív munkás keresők csak relatív többségben vannak, azaz arányuk 25 és 50 százalék közé esik /vegyes munkás jellegű körzetek, illetve települések/
3. réteg: a szövetkezeti parasztsághoz tartozó aktív keresők vannak abszolút vagy relatív többségben /paraszti jellegű körzetek, illetve települések/
4. réteg: az egyéb /szellemi, kisárutermelő, kiskereskedő/ aktív keresők vannak abszolút vagy relatív többségben
5. réteg: az inaktív keresők vannak abszolút vagy relatív többségben /inaktív jellegű körzetek, illetve települések/

A külterületi körzeteket részben viszonylag csekély számuk, részben speciális népesség összetételük miatt csak az alábbi két rétegbe soroltuk:

6. réteg: a paraszti és inaktív keresők többen vannak, mint a munkás és egyéb keyesők

7. réteg: a munkás és egyéb keresők vannak többségben.

A körzetek rétegzésével kapcsolatban még lényeges megjegyezni, hogy azokat előzőleg három nagyságkategóriába soroltuk:

a/ a körzet magánháztartásokban élő lakónépessége nagyobb 100-nál

b/ a körzet magánháztartásokban élő lakónépessége nagyobb 10-nél, de nem nagyobb 100-nál

c/ a körzet magánháztartásokban élő lakónépessége nem nagyobb 10-nél.

Az a/ kategóriába eső körzetek voltak az ún. normál körzetek, ezek mind a rétegzés, mind a kiválasztás során önálló egységként szerepeltek. A b/ kategóriába került körzetek nem képeztek önálló egységet, hanem vagy egy földrajzilag közelfekvő normál körzethez csatoltuk, vagy több ilyen kis körzet összevontan képezett egy normál körzetet. A csatolás, illetve összevonás még természetesen a rétegzés előtt történt, azaz az összevont körzetek együttes aktív és inaktív kereső összetételük alapján kerültek a fenti hét réteg valamelyikébe. A c/ kategóriába csak kevés körzet került, ezek össznépessége is elenyésző, így gyakorlati okok miatt ilyen körzetek nem kerültek kiválasztásra.

### 2.3 A TELEPÜLÉSEK KIVÁLASZTÁSA ÉS A TELEPÜLÉS-MINTA ELLENŐRZÉSE

Mint említettük, a kedvező tapasztalatok alapján az új ELAR mintánál is a nagysággal arányos valószínűségű, visszatevés nélküli kiválasztási eljárást alkalmaztuk mind a települések, mind azon belül a körzetek kiválasztására. Ehhez olyan táblák álltak rendelkezésünkre, amelyek megyénként, azon belül az 1.-6. település nagyságkategóriákra külön-külön tartalmazták a településeket, azok 1980. évi népszámlálásnál regisztrált lakásszámát, a kumulált /folyamatosan

összegzett/ lakásszámokat, valamint a települések jelleg szerinti rétegazonosítóját. /A 7.-9. nagyságkategóriákban, azaz a 10 ezernél népesebb települések közül mint jeleztük, nem történt kiválasztás, ezek mind szerepeltek az arányos mintában./ Az egyes nagyságkategóriákon belül a települések /az akkor még meglévő/ járásokként /városkörnyékeként/, azon belül abc szerint voltak sorbarendezve, így a követett kiválasztási eljárás impliciten egy megyén belüli területi rétegzést is magába foglalt.

Az egyes nagyságkategóriákból kiválasztandó települések számának meghatározása, illetve a kiválasztás végrehajtása a következőképpen történt:

- a/ minden nagyságcsoportban meghatároztuk a normál /100-nál nagyobb lakónépességű, illetve ilyenre egyesített/ körzetek számát, majd ennek 15 %-át. Jelölje ezeket rendre  $r_1, \dots, r_6$ .
- b/ Az 1 500-nál kisebb lélekszámú községek esetén /1.-3. nagyságcsoportok/ a kiválasztandó községek száma közelítőleg  $0,15 \cdot N_i$  volt, ahol  $N_i$  az  $i$ -edik nagyságcsoportban a községek száma. Ha ugyanitt  $L_i$  jelöli az összes lakásszámot, akkor a kiválasztásnál a lépésközt  $k_i = L_i / 0,15 N_i$  adta. Meghatározva egy  $k_i$ -nél nem nagyobb véletlen kezdőszámot,  $Z_i$ -t, azok a települések kerültek kiválasztásra, amelyek kumulált lakásszáma éppen nem haladta meg a  $Z_i, Z_i + k_i, Z_i + 2k_i, \dots$  értékeket. A kiválasztás ténylegesen addig folytatódott, amíg a kiválasztott községek normál körzeteinek száma el nem érte  $r_i$ -t. Mivel általában nem lehetett biztosítani - különösen az 1 000-1 500 lélekszámú községeket tartalmazó 3. nagyságkategória esetén -, hogy a kiválasztott községek együttes körzetszáma pontosan  $r_i$  legyen, a + vagy - eltérést a következő nagyságcsoportnál vettük figyelembe, hogy összességében kiegyenlítődjenek az eltérések.
- c/ A 4.-6. nagyságkategóriák esetén a kiválasztandó községek számát  $r_i / 6$  adta /keresítve a legközelebbi egész számra/, hiszen itt a mintavételi terv minden településről 6 normál körzet kiválasztását írta elő. A lépésköz így  $6L_i / r_i$  volt, egyébként a kiválasztás végrehajtása ugyanúgy történt,

mint az 1.-3. nagyságkategóriában. A + vagy - eltérést a kiválasztott települések hatszorosa és  $r_1$  között itt is hozzáadtuk a következő nagyságcsoport összkörzetszáma 15 %-ához, azaz  $r_{i+1}$ -hez.

Mint a fenti leírásból látható, a települések jelleg szerinti rétegei a kiválasztás során nem lettek direkt módon rétegenként figyelembe véve. Utólag azonban minden nagyságcsoporton belül meghatároztuk a j-edik réteghez tartozó községek számát,  $N_{ij}$ -t, illetve e községek normál körzeteinek 15 %-át,  $r_{ij}$ -t, és ellenőriztük, hogy a j-edik rétegből kiválasztott községek száma közelítőleg megegyezik-e  $0,15 N_{ij}$ -vel, illetve  $r_{ij}/6$ -tal. Jelentősebb eltérés esetén véletlen településcserével biztosítottuk, hogy a jelleg szerinti rétegek is megfelelően legyenek képviselve a településmintában.

A leirt mintavételi eljárással végeredményben 561 10 ezernél kisebb település került az arányos mintába, ezek közül 9 város volt, a többi község. A 137 10 ezernél nagyobb települést mind tartalmazta az arányos 15 %-os minta, így az összesen 698 településre terjedt ki, Budapesten és az öt megyei városon kívül 101 városra és 591 községre.

A kiválasztás végrehajtása után részletesen ellenőriztük, hogy a településminta hogyan reprezentálja a 10 ezernél kisebb települések összességét. E célból a népszámlálás során összeírt 30 ismérvre becslést számítottunk gépi úton a mintából a visszatevés nélküli, nagysággal arányos valószínűségű rétegezett csoportos mintavételnek megfelelő becslő formula alapján megyénként és ezen belül külön az 1 500-nál kisebb és az 1 500-10 000 lélekszámú településekre. Kiszámítottuk továbbá e becslések 95 %-os valószínűségi szintű megbízhatósági határait és minden egyes ismérvnél megvizsgáltuk, hogy a megfelelő tényleges népszámlálási adatok e megbízhatósági határok közé esnek-e vagy sem. Az esetek túlnyomó többségében a megbízhatósági intervallumok tartalmazták a teljeskörű adatokat. Néhány jelentősebbnek ítélt eltérés esetén egy-két település rétegen belüli véletlen cseréjével biztosítottuk, hogy ezek a becslések is a kívánt irányban javuljanak. Az ellenőrzésbe bevont 30 népszámlálási ismerv az alábbi:

lakónépesség

lakásszám

férfiak száma összesen és négy korcsoportban

nők száma összesen és négy korcsoportban

aktív keresők száma nemenként és összesen  
 aktív keresők száma 3 népgazdasági ág szerinti csoportban  
 aktív keresők száma 4 alapvető osztály és rétegcsoportban  
 inaktív keresők száma nemenként  
 eltartottak száma nemenként  
 lakások száma az építési év szerinti 3 kategóriában  
 háztartások száma.

Ha  $n_i$  jelöli az  $i$ -edik nagyságcsoportból kiválasztott települések számát  $i=1,2,3$ , illetve  $4,5,6$ , akkor az  $i$ -edik nagyságcsoportba tartozó  $j$ -edik település kiválasztási valószínűsége a leírt mintavételi eljárásnál:

$$P_{ij} = \frac{L_{ij}}{L_i} n_i,$$

ahol  $L_{ij}$  az adott község lakásainak száma. Így valamely ellenőrzendő ismerv becslése a településmintából:

$$\hat{X} = \sum_{i=1}^3 \hat{X}_i = \sum_{i=1}^3 \sum_{j=1}^{n_i} x_{ij}/P_{ij} = \sum_{i=1}^3 \frac{L_i}{n_i} \sum_{j=1}^{n_i} x_{ij}/L_{ij}$$

ahol  $x_{ij}$  jelöli a mintába került  $j$ -edik településen az  $X$  ismérvnek a népszámlálásnál regisztrált értékét. Az 1 500-10 000 lakónépességű települések esetén természetesen  $i$  nem 1-től 3-ig, hanem 4-től 6-ig fut.

Az  $\hat{X}$  becslés szórását nem a mintából kellett becsülni, hanem a teljes alapsokaságból lehetett számítani a

$$\sigma_{\hat{X}} = \sqrt{\sum_{i=1}^3 \left\{ \frac{L_i}{n_i} \sum_{j=1}^{n_i} L_{ij} / \bar{x}_{ij} - \bar{x}_i / 2 - \frac{n_i - 1}{n_i} \sum_{j=1}^{n_i} L_{ij}^2 / \bar{x}_{ij} - \bar{x}_i / 2 \right\}}$$

formula<sup>2/</sup> alapján, ahol

$$\bar{x}_{ij} = x_{ij}/L_{ij} \quad \text{és} \quad \bar{x}_i = x_i/L_i$$

2/ A szórásformulána vonatkozólag l. pl. H.O.Hartley: Systematic sampling with unequal probability and without replacement. JASA 61 /1966/ 739-748 pp.



nál a gépi listák a körzeteket rétegenként tartalmazták, így a szisztematikus kiválasztási eljárás automatikusan biztosította az egyes rétegek arányos képviselőjét a mintában.

A 10 ezernél népesebb településeknél is ellenőriztük és esetenként a körzet rétegen belüli cseréjével biztosítottuk, hogy régi ELAR körzetek ne szerepeljenek az új mintában.

Az 1 500-nál népesebb településeknél a számlálókörzeti mintára vonatkozólag is sor került olyan ellenőrzésre, mint a település minta esetén. Az ellenőrzés megynként történt külön a 4-6., illetve 7-8. nagyságcsoportu településekre, illetve a 25 ezernél népesebb városokra egyenként. Az ellenőrzés során a körzetmintából a kiválasztási eljárásnak /egy-, illetve kétlépcsős rétegzett, csoportos, nagysággal arányos valószínűségű, visszatevés nélküli szisztematikus kiválasztás/ megfelelő becslések készültek a már említett 30 népszámlálási ismérvre, meghatároztuk e becslések alapsokaságbeli - tehát nem a mintából becsült - szórását, 95 %-os megbízhatósági szintű konfidencia határait, majd egyenként ellenőriztük, hogy a tényleges népszámlálási értékek belül vannak-e a számított konfidencia intervallumon.

a/ A 4-6. nagyságcsoportu települések esetén az  $i$ -edik nagyságcsoport  $j$ -edik községe  $k$ -adik körzetének kiválasztási valószínűsége:

$$P_{ijk} = P_{ij} \cdot P_{k|ij} = n_i \frac{L_{ij}}{L_i} \cdot 6 \frac{L_{ijk}}{L_{ij}} = 6 n_i \frac{L_{ijk}}{L_i} \quad \begin{matrix} i=4,5,6 \\ j=1,\dots,N_i \\ k=1,\dots,K_{ij}^1 \end{matrix}$$

s így valamely  $X$  ismerv /értékösszeg/ becslése a körzetmintából az  $i$ -edik nagyságcsoportban

$$\hat{x}_i = \sum_{j=1}^{n_i} \sum_{k=1}^6 \frac{x_{ijk}}{P_{ijk}} = \frac{L_i}{6n_i} \sum_{j=1}^{n_i} \sum_{k=1}^6 \frac{x_{ijk}}{L_{ijk}} = \frac{L_i}{6n_i} \sum_{j=1}^{n_i} \sum_{k=1}^6 \bar{x}_{ijk}$$

és a 4-6. nagyságcsoportu települések összességére

$$\hat{x} = \sum_{i=4}^6 \hat{x}_i = \frac{1}{6} \sum_{i=4}^6 \frac{L_i}{n_i} \sum_{j=1}^{n_i} \sum_{k=1}^6 \bar{x}_{ijk}$$

Az előzőekben  $L_{ijk}$  a lakások száma az  $i$ -edik nagyságcsoporthoz  $j$ -edik községre  $k$ -adik körzetében,  $L_{ij}$  ugyanez a  $j$ -edik község,  $L_i$  pedig az egész nagyságcsoporthoz egészére.  $\bar{x}_{ijk}$  az  $X$  ismérv egy lakásra jutó értéke az  $\{i, j, k\}$  körzetben. Mivel a 4-6. nagyságcsoporthoz települések mindegyikéből 6 körzet került kiválasztásra, a becslésnél az összegezés  $k$  szerint 1-től 6-ig megy.

$\hat{X}$  szórásnégyzete az  $\hat{X}_i$  becslések szórásnégyzetének összege, hiszen a minta kiválasztása az egyes nagyságcsoporthoz egymástól függetlenül történt, azaz az  $\hat{X}_i$  becslések függetlenek:

$$\sigma_{\hat{X}}^2 = \sum_{i=4}^6 \sigma_{X_i}^2 = \sum_{i=4}^6 \frac{1}{n_i} \sum_{j=1}^{N_i} \left\{ L_{ij} \left[ \frac{L_i - 1}{L_{ij}} \right] \frac{(\bar{x}_{ij} - \bar{X}_i)^2}{L_{ij}} \right. \\ \left. + \frac{L_i}{6} \sum_{k=1}^{K_{ij}} \frac{L_{ijk}}{L_{ij}} \frac{(\bar{x}_{ijk} - \bar{x}_{ij})^2}{L_{ijk}} \right\}$$

A 95 %-os megbízhatósági szintű konfidencia határok pedig:

$$\hat{X} \pm 1,96 \sigma_{\hat{X}}$$

b/ A 7. és 8. nagyságcsoporthoz települések mind szerepelnek az arányos mintában. Az  $i$ -edik nagyságcsoporthoz  $j$ -edik település  $k$ -adik körzetének kiválasztási valószínűsége:

$$P_{ijk} = 0,15 K_i \frac{L_{ijk}}{L_i} \quad i=7,8; \quad j=1, \dots, N_i, \quad k=1, \dots, K_{ij},$$

ahol  $K_i = \sum_{j=1}^{N_i} K_{ij}$  az  $i$ -edik nagyságcsoporthoz összes körzetének a száma. Ennek megfelelően valamely  $X$  értékösszeg becslése a körzetmintából az  $i$ -edik nagyságcsoporthoz

$$\hat{X}_i = \frac{L_i}{0,15 K_i} \sum_{j=1}^{N_i} \sum_{k=1}^{K_{ij}} x_{ijk} / L_{ijk} = \frac{L_i}{0,15 K_i} \sum_{j=1}^{N_i} \sum_{k=1}^{K_{ij}} \bar{x}_{ijk}$$

ahol  $k_{ij}$  jelöli a  $j$ -edik településből a mintába került körzetek számát.

A két nagyságcsoporthoz egészére a becslés

$$\hat{X} = \hat{X}_7 + \hat{X}_8$$

Mivel itt nem történt település kiválasztás, a szórásnégyzet formula kissé különbözik az a/-ban szereplőtől. Itt az alapsokaságból számított szórásnégyzet:

$$\sigma_{\hat{x}}^2 = \sigma_{\hat{x}_7}^2 + \sigma_{\hat{x}_8}^2 = \sum_{i=7}^8 \frac{L_i}{0,15K_i}$$

$$\left\{ \sum_{j=1}^{N_i} \sum_{k=1}^{K_{ij}} \frac{L_{ijk}}{L_i} \left[ L_i - 1/0,15K_i - 1/L_{ijk} \right] / \bar{x}_{ijk} - \bar{x}_{ij} / 2 \right\}$$

Ugyanez a becslő és szórásnégyzet formula alkalmazható Budapest és Miskolc esetében, ahol kerületenként történt a kiválasztás. E két város esetében természetesen az i index és az erre való összegezés elmarad, j Budapest esetén 1-től 22-ig, Miskolcnál pedig 1-től 3-ig megy.

c/ A 9-es nagyságcsoportú városok is mind szerepelnek a mintában és ezekre a körzetminta ellenőrzése városonként történt. Egy adott város k-adik körzetének kiválasztási valószínűsége

$$P_k = 0,15K \frac{L_k}{L}$$

ahol K, illetve L a városban a körzetek, illetve lakások száma. Így valamely X értékösszeg becslése a körzetmintából:

$$\hat{x} = \frac{L}{0,15K} \frac{x_k}{L_k} = \frac{L}{0,15K} \bar{x}_k$$

és e becslés megbízhatósági határai 95 %-os valószínűségi szinten:

$$\hat{x} \pm 1,96 \sqrt{\frac{1}{0,15K} \sum_{k=1}^K L_k \left[ L - 1/0,15K - 1/L_k \right] / \bar{x}_k - \bar{x} / 2}$$

ahol  $\bar{x}$  a vizsgált ismerv egy lakásra jutó értéke az adott városban.

A körzetminta ellenőrzése megnyugtató eredményt adott. Az esetek túlnyomó többségében a valódi és becsült értékek eltérése jóval kisebb volt a megengedett hibahatárnál. Csak néhány esetben fordult elő, hogy valamely egységben több vizsgált ismérv valódi /népszámláláskor regisztrált/ értéke kívül esett a számított konfidencia határokon. Ilyenkor megfelelő összetételű körzetek véletlen cseréjével javítottuk az arányos minta reprezentativitását.

## 2.5 A MINTA KIEGÉSZÍTÉSE ÚJ LAKÓTELEPI KÖRZETEKSEL

A fentiek szerint kialakított és ellenőrzött arányos körzetminta természetesen csak a népszámlálás idején meglévő körzeteket, illetve lakásállományt reprezentálta. Az új ELAR mintának azonban induláskor az 1983. január 1-i lakásállományt és később is mindig az aktuális lakásállományt kellett, illetve kell reprezentálnia. A 15 %-os arányos körzetminta elég nagy minta ahhoz, hogy joggal feltehessük, hogy a kiválasztott körzetek jól reprezentálják az 1980. január 1. és 1982. december 31. közötti ún. szórt, azaz nem lakótelepi jellegű építkezéseket /és bontásokat/. /A mintakörzetek cimanyagának naprakész tételéről 1. később a 3. pontot./ A lakótelepi építkezéseket azonban az 1980. évi körzetekből kiválasztott minta nem tudja reprezentálni, hiszen ezen építkezések során új körzetek keletkeztek. Ezért szükség volt a mintát kiegészíteni az 1980. január 1. után épült és 1982. december végéig átadásra tervezett lakótelepi lakásokból képzett virtuális körzetek 15 %-ával. A körzetek kiválasztása ez esetben is nagysággal arányos valószínűséggel történt. Biztosítottuk, hogy a minta ilyen szempontból is arányos legyen, azaz Budapestről, illetve az egyes megyékből, azon belül a különböző nagyságu településekről a jelzett időszakban épített lakótelepi körzetek számával arányosan kerüljenek be ilyen új körzetek a mintába.

## 2.6 A MINTA ARÁNYOSSÁGÁNAK ÉS A LAKÓTELEPI REPREZENTÁCIÓNAK ELLENŐRZÉSE

A települések és körzetek kiválasztási módszeréből adódóan egy konkrét felvétel esetén akkor van egy rétegen belül minden lakásnak egyforma esélye a mintába kerülésre, ha körzetenként azonos

számu lakás kerül kiválasztásra. A megyénként és rétegenként 15 %-os körzet minta viszont csak akkor lesz valóban arányos körzetenként azonos számu lakás kiválasztása mellett, ha a körzetek átlagos lakásszáma különböző rétegekben megegyezik. Az arányos minta véglegesítése előtt ezért még ellenőrizni kellett a minta pontos arányosságát oly módon, hogy körzetenként 10 lakás kiválasztását feltételezve megnéztük, hogy az így adódó lakásszámok megyénként, illetve rétegenként a lakásokra nézve azonos kiválasztási arányt eredményeznek-e. Olyan rétegeknél, ahol az egy körzetre jutó átlagos lakásszám jelentősen meghaladta az országos átlagot, vagy elmaradt attól, s emiatt számottevő eltérés mutatkozott az arányos mintától, egy-egy körzet elhagyásával, illetve pótlólagos kiválasztásával biztosítottuk, hogy minden olyan konkrét minta, amely a körzet mintából körzetenként azonos számu lakás kiválasztásával áll elő, valóban arányos legyen, azaz a lakásokra vonatkozó kiválasztási arány minden rétegben megegyezzen.

Még egy további szempontból is ellenőriztük a mintát, mégpedig a lakótelepi lakások reprezentációja szempontjából. A körzetek kiválasztása során az egyes foglalkozási rétegeken belül nem különböztettük meg a lakótelepi és nem lakótelepi körzeteket. Ugyanakkor ismeretes, hogy a lakótelepeken lakó háztartások összetétele, jövedelmi, fogyasztási, lakásviszonyai stb. sok tekintetben eltérnek a nem lakótelepi háztartásokétól. A tervezett ELAR felvételek adatainak megbízhatósága, általánosíthatósága szempontjából ezért lényeges volt, hogy az arányos mintában - ismét körzetenként azonos számu lakás kiválasztását feltételezve - a lakótelepi lakásokra vonatkozó kiválasztási arány azonos legyen a nem lakótelepi lakásokéval. Ez az ellenőrzés természetesen nem az egész mintát, inkább csak a nagyobb településeket érintette, ahol egyáltalán vannak lakótelepek. A rétegek tulajdonképpénységében már az eredeti minta is megfelelő arányban reprezentálta a lakótelepi lakásokat, néhány esetben körzetcserével biztosítottuk a megfelelő reprezentációt.

Az ily módon sokoldaluan ellenőrzött és valóban arányos körzetminta szolgált alapul a mintavétel második fázisában a nem arányos, működő ELAR minta kiválasztásához.

### 3. A MŰKÖDŐ NEM ARÁNYOS MINTA KIVÁLASZTÁSA

#### 3.1 A NEM ARÁNYOS MINTA KIALAKÍTÁSA

Mint már erről szó volt, az arányos mintától való eltérést egyrészt egyes kis sulyu társadalmi rétegek jobb reprezentációja, másrészt a minta több településre történő kiterjesztése révén elérhető nagyobb megbízhatóság indokolta. Az arányos minta nem arányos csökkentése a 10 ezernél nem népesebb települések esetén települések elhagyása révén valósult meg, mégpedig az első célnak megfelelően az alábbi módon:

- az arányos minta jelleg szerint 3. rétegbe tartozó, azaz paraszti jellegű települései lényegében mind szerepelnek a végső mintában is;
- az 1., 2. és 4. rétegbe tartozó, azaz nem paraszti aktív jellegű településeknek nagyjából a fele;
- az 5. rétegbe tartozó, azaz előregedett, inaktív jellegű településeknek kb. negyedrésze maradt meg a nem arányos mintában.

A nem arányos mintában maradó települések kiválasztása az arányos minta települései közül már egyenlő valószínűséggel történt megynként és az egyes nagyságkategóriákon belül külön-külön. Mivel az arányos minta nem paraszti aktív, illetve inaktív jellegű településeinek száma egy-egy nagyságkategórián belül esetenként nem volt páros, illetve négy többszöröse, ezért a vonatkozó kiválasztási arányok nem mindig pontosan 50, illetve 25 % voltak, a mintanagyságot behatároló költségkeretek miatt ezektől inkább lefelé tértek el.

További eltérést okozhatott egyes esetekben a tervezett kiválasztási arányoktól az 1-3. nagyság csoportokban az a tény, hogy itt a maradó települések kiválasztása a 2-3 kis községből álló mintavételi egységek kiválasztásával történt, és ezek a földrajzilag szomszédos és megfelelő közlekedési összeköttetéssel rendelkező, egy mintavételi egységet alkotó községek jelleg szerint nem feltétlenül tartoztak azonos rétegbe.

Körzet kiválasztás, illetve elhagyás az 1-6. nagyságcsoportokban ebben a fázisban nem történt, a működő végső mintában szereplő 1-3. nagyságcsoportú községek teljes cimanyagukkal, a 4-6. nagyságkategóriába tartozó települések pedig az első fázisban kiválasztott 6 körzet cimanyagával kerültek be a nem arányos körzetmintába.

A 10-25 ezer lélekszámú települések esetén a nem arányos kiválasztás nem a települések, hanem a körzetek szintjén valósult meg. A jelleg szerint különböző rétegekbe tartozó körzetekre hasonló kiválasztási arányokat alkalmaztunk, mint a 10 ezernél kisebb településekre. A kiválasztandó települések számát - a mintavétel második fázisában ugyanis szükségszerűen a 10-25 ezer lakosu települések közül is ki kellett választani bizonyos számot - részben a mintában maradó különböző rétegű számlálókörzetek száma, részben pedig az a követelmény határozta meg, hogy a 7-es nagyságcsoportú települések mindegyikében 6, a 8-as nagyságkategóriában pedig településenként 8 körzet maradjon a nem arányos mintában. Így ha  $k_p$ ,  $k_a$ , illetve  $k_i$  jelöli az arányos mintában a különböző kiválasztási arányu /paraszti, nem paraszti aktív, illetve inaktív jellegű/ rétegekbe tartozó körzetek számát, akkor a kiválasztandó települések számát az

$$n_7 = \frac{k_p + k_a / 2 + k_i / 4}{6} \quad \text{és} \quad n_8 = \frac{k_p + k_a / 2 + k_i / 4}{8}$$

formulák adják egészre kerekítve. A települések kiválasztása itt is nagysággal arányos valószínűséggel történt.

A 25 ezernél népesebb városok a nem arányos mintában is mind szerepelnek, azaz önreprezentálók. A minta nagyobb terítettsége - több településre való kiterjedése - érdekében e városokban az arányos minta körzeteiből nagyobb arányban történt elhagyás, mint más nagyságkategóriákban. Különösen vonatkozik ez a megyei városokra és Budapestre. Így végeredményben e nagyságkategóriában az arányos minta nem paraszti aktív jellegű körzeteinek csak kevesebb mint 40 %-a, az inaktív jellegűeknek pedig kevesebb mint 20 %-a maradt meg a nem arányos ELAR mintában. A maradó körzetek kiválasztása egy-egy rétegen belül egyenlő valószínűséggel történt. Összességében a 25 ezernél népesebb nem megyei városokban az arányos minta körzeteinek 42 %-a, a megyei városokban 37 %-a, Budapesten pedig 31 %-a maradt meg a nem arányos mintában. Az eltérő arányokba természetesen

az is belejátszott, hogy a városok e három kategóriájában különböző volt a paraszti, a nem paraszti aktív, illetve inaktív jellegű körzetek súlya mind az alapsokaságban, mind ebből kifolyólag az arányos mintában.

Végeredményben az új, nem arányos ELAR minta 405 településen 2 630 körzetet foglal magába, beleértve az 1980-82 években épült új lakótelepi körzetek közül kiválasztott körzeteket is.

### 3.2 AZ ARÁNYOS ÉS NEM ARÁNYOS MINTA ÖSSZETÉTELE

Célszerűnek látszik részletesen is bemutatni az arányos, illetve a nem arányos ELAR minta összetételét település nagyság-kategóriák szerint, jelezve egyúttal a körzetekre vonatkozó eltérő kiválasztási arányokat.

1. táblázat

A 15 %-os arányos és a működő ELAR minta összetétele település kategóriák szerint

Település kategória	A mintába kiválasztott				% -os aránya
	települések		számlálókörzetek		
	száma				
	15 %-os arányos minta	ELAR minta	15 %-os arányos minta	ELAR minta	
<b>Községek</b>					
1 000-nél kevesebb lakossal	217	127	401	266	10,0
1 000- 1 499 lakossal	66	33	280	136	7,3
1 500- 2 999 lakossal	116	66	660	389	8,8
3 000- 4 999 lakossal	72	40	430	240	8,4
5 000- 9 999 lakossal	81	38	486	228	7,0
10 000-14 999 lakossal	30	16	189	98	7,8
15 000 és több lakossal	9	5	81	38	7,0
Községek együtt	591	325	2 527	1 395	8,3
<b>Városok<sup>a/</sup></b>					
10 000-nél kevesebb lakossal	9	5	55	31	8,5
10 000-14 999 lakossal	21	8	125	58	7,0
15 000-24 999 lakossal	28	18	288	130	6,8
25 000 és több lakossal	43	43	1 094	462	6,3
Városok együtt	101	74	1 562	681	6,5
Megyei városok	5	5	483	178	5,5
Budapest	1	1	1 218	376	4,6
Ország összesen	698	405	5 790	2 630	6,8

a/ Az új minta indulásakor, 1983. január 1-én érvényes közigazgatási beosztásnak megfelelően.

Mint látható, a körzetekre vonatkozó kiválasztási arányban több mint kétszeres a különbség Budapest és a legkisebb község kategória között. Ha azonban a községek, illetve Budapesten a körzetek jelleg szerinti rétegei külön nézzük a kiválasztási arányokat, akkor a szélső esetekben ötszörös különbségeket is találhatunk. Ha a nagyobb városokban nem alkalmaztunk volna kisebb kiválasztási arányt, mint a községekben, akkor azonos mintanagyság - körzetszám - mellett a minta kb. 70-nel kevesebb községet foglalt volna magába, azaz a községeket tekintve több mint 20 %-kal kevesebb elsődleges mintavételi egységet. Ez feltétlenül negatív hatással lett volna az ELAR felvételek községekre, illetve kisebb területi egységekre vonatkozó adatainak megbízhatóságára.

### 3.3 KÉT RÉSZ-MINTA KIALAKÍTÁSA

Bár az új ELAR minta az előzőtől eltérően a településeket tekintve egységes, integrált minta, célszerűnek látszott most is a körzetekből két független rész-mintát képezni, egyiket a háztartási költségvetési felvétel és a hozzá kapcsolódó egyéb felvételek, a másikat a társadalmi-demográfiai alapadat felvétel /régebben munkaerő-felvétel/ és a kapcsolódó egészségügyi, társadalomstatistikai, stb. felvételek céljára. A két rész-minta kialakítását nemcsak az indokolta, hogy a két alapfelvétel az 1983-as kezdőévet leszámítva időben is elválik - páros években kerül sor a társadalmi-demográfiai alapadat felvételre, páratlan években a háztartási költségvetési felvételekre -, s így bizonyos munka megtakarítás érhető el, pl. egy-egy évben csak a körzetek felét kell bejárni a címlista naprakésszé tétel érdekében, hanem elsősorban az, hogy a többi ELAR felvételtől eltérően a háztartási költségvetési felvételnél a kiválasztás alapegysége nem a lakás, hanem a háztartás. Körzetek szintjén is egységes minta esetén ez zavarokat okozott volna a címek, illetve háztartások kiválasztásánál.

Megjegyzendő azonban, hogy továbbra is vannak olyan felvételek - pl. az öt évenként végrehajtásra kerülő jövedelmi felvételek, az 1983. évi mobilitási felvétel vagy az 1984. évi ifjúsági felvétel -, amelyek mindkét rész-mintára, azaz az ELAR minta minden körzetére kiterjednek.

A két rész-minta kialakításához a legkisebb - egy körzetből álló - községek kivételével minden településből eleve páros számú körzet került a mintába. Így a két - "H", illetve "M" jelű<sup>3/</sup> - rész-minta a körzetek településenként két véletlen csoportra bontásával adódott, gondoskodva arról, hogy a részmintákban a körzetek összetétele foglalkozási réteg, külterület-belterület, illetve lakótelepi - nem lakótelepi körzet szerint megyénként és település nagyságcsoportonként teljesen azonos legyen. Az egy számlálókörzetből álló kis községek váltakozva kerültek az egyik vagy másik részmintába. Végeredményben az "M" rész-minta 390 településen 1 314 körzetet, a "H" rész-minta pedig 382 településen 1 316 körzetet tartalmaz.

A két rész-minta kialakításával lezárult az új ELAR minta, mint az ELAR keretében 10 éven keresztül végrehajtásra kerülő felvételek mintáihoz szolgáló mintavételi keret kiválasztása. Az egyes konkrét felvételek céljára a lakások, illetve háztartások kiválasztása az "M" vagy a "H" rész-minta - vagy a teljes minta - körzeteiből körzetenként azonos számú lakás /háztartás/ véletlen /egyenlő valószínűségű/ kijelölésével történik. Így pl. a társadalmi-demográfiai alapadat felvételhez az "M" részmintából körzetenként 12 lakás, a háztartási költségvetési felvételhez a "H" rész-minta körzeteiből 9 háztartás, az 1983. évi jövedelmi felvételhez pedig az összes körzetből 6 lakás kerül, illetve került kiválasztásra szisztematikus mintavétellel.

### 3.4 SÚLYOZÁSI ALGORITMUS

Mivel az új ELAR minta nem arányos elosztású, az egyes felvételek adatait csak súlyozással - teljeskörűsítéssel - lehet feldolgozni. A mintavételi elméletből ismeretes, hogy nem egyenlő valószínűségű kiválasztás esetén az egyes mintaelemeket a kiválasztási valószínűség reciprokával szorozva lehet összesíteni. Egy konkrét

3/ "H": háztartási költségvetési felvétel céljait szolgáló rész-minta. - "M": munkaerő /társadalmi-demográfiai alapadat/ felvétel és egyéb kapcsolódó felvételek céljait szolgáló rész-minta.

mintánál az egyes rétegekben a kiválasztási valószínűséget a mintába került lakások /háztartások/ és a rétegben a vizsgált alapsokasághoz tartozó összes lakás /háztartás/ hányadosa adja. A súlyozási algoritmushoz országosan mintegy 270 rétegcsoport kialakítására került sor oly módon, hogy a CLUSTERS hibaszámítási program alkalmazhatósága érdekében ezek mindegyike mindkét részmintából legalább két elsődleges mintavételi egységet /települést, illetve körzetet/ tartalmazzon. A 270 különböző rétegcsoport nem feltétlenül jelent ugyanennyi különböző kiválasztási arányt, de ahhoz, hogy az adatokat megyénként, illetve város-község bontásban is fel lehessen dolgozni, ezeket a rétegeket megyénként kellett kialakítani, úgy, hogy a városok és községek feltétlenül különböző rétegcsoportokba kerüljenek. Így végeredményben Budapesten 50, a megyék többségében 11-17, a kisebb megyékben 6-9 rétegcsoport kialakítására került sor.

Olyan felvételeknél, ahol a lakás a végső kiválasztási egység, az egységes súlyozási algoritmus minden rétegcsoporthoz egy súlyszámot - felszorzó faktort - rendel hozzá, amely nem más, mint a szöbanforgó réteg építési és bontási statisztikák alapján továbbvezetett teljeskörű lakott lakásszámának és az adott felvételnél a rétegből ténylegesen megfigyelt lakások számának hányadosa, azaz a kiválasztási valószínűség reciproka. Az építési és bontási statisztikák alapján természetesen csak az összes lakásszámot lehet évről-évre továbbvezetve megállapítani, a lakott lakások számát a mintából kell becsülni a mintában egy-egy rétegcsoportban talált üres /nem lakott/ lakások aránya alapján. Így, ha  $p_o /i/$  jelöli az  $i$ -edik rétegcsoportban az üres lakások arányát a mintában,  $L_i$  a továbbvezetett teljes lakásszámot és  $l_i$  az adott felvételen ténylegesen szereplő lakások számát, akkor az  $i$ -edik rétegcsoportban alkalmazandó súlyszámot az

$$s_i = \frac{L_i [ 1 - p_o /i/ ]}{l_i}$$

formula adja egész számra kerekítve. A súlyszámok nagyságrendjének érzékeltetésére megemlítjük, hogy pl. az 1986. évi társadalmi-demográfiai alapadat felvétel esetén a súlyszámok többsége 200-400 között van, de 15 esetben 100-nál kisebb a súlyszám és kb. ugyanennyi esetben nagyobb mint 600. A nagyobb súlyszámok többsége budapesti, illetve nagyvárosi rétegcsoportokhoz tartozik.

Olyan felvételeknél, amelyek nem a magánháztartásokban élő teljes népességre, hanem annak csak egy meghatározott részére vonatkoznak, de a végső kiválasztási egység továbbra is a lakás, tulajdonképpen ugyanez a formula alkalmazható a súlyszám meghatározására, csak az üres lakás fogalmát kell módosítani, oly módon, hogy üres lakásnak minősül minden olyan lakás, ahol nem lakik az alapsokasághoz tartozó személy /háztartás/. Az 1984. évi ifjusági felvétel pl. a 15-35 éves személyekre vonatkozott, így e felvételnél üres lakásnak minősült minden olyan lakás, ahol nem lakott a vizsgált korosztályba tartozó személy. Természetesen az így definiált üres lakások arányát is rétegcsoportonként a mintából kellett becsülni. A rétegcsoportok többségében a kiválasztott lakások száma elég nagy ahhoz, hogy a szűkebb vagy tágabb értelemben vett üres lakások arányát elég megbízhatóan lehessen becsülni, de természetesen ez is, mint minden mintából kapott becslés, bizonyos véletlen hibát tartalmaz. A súlyszámokban jelentkező kisebb hibák azonban a végső eredményeket - részben a rétegcsoportok nagy száma, részben a vizsgált jelenségek többségének rétegcsoportok közötti mérsékelt szóródása miatt - csak csekély mértékben befolyásolják.

A háztartási költségvetési felvételnél nem a lakás, hanem a háztartás a végső kiválasztási egység. E mellett ez a felvétel sem terjed ki a teljes népességre, olyan háztartások, amelyekben önálló foglalkozási viszonyú kereső vagy külföldi személy él, nem tartoznak a megfigyelni kívánt alapsokasághoz. Ez esetben a megfelelő súlyszám megállapításához nemcsak a tágabb értelemben vett üres lakások arányát kellett a mintából becsülni, hanem az egy lakásra jutó háztartások átlagos számát is, hiszen ebben az esetben a súlyszám számlálója a rétegcsoportban az alapsokasághoz tartozó háztartások teljeskörű száma.

### 3.5 AZ ÚJ ELAR MINTA ELLENŐRZÉSE

Azon túlmenően, hogy a nem arányos minta kiválasztása során folyamatosan ellenőriztük, hogy az ELAR minta összetétele egy-egy azonos kiválasztási arányú rétegen belül hűen tükrözi-e az arányos mintáét - s így egyuttal az alapsokaságét -, a minta tényleges felhasználása előtt teszt számítások is készültek a minta reprezentativi-

tása és egyuttal a súlyozási algoritmus ellenőrzésére. E célból az "M" részminta népszámlálási körzeteiből gépi véletlen kiválasztási eljárással körzetenként 12 cím került kijelölésre, és e lakásokban lakó személyek, illetve háztartások népszámlálási adatai alapján becslés - teljeskörűsített adat - készült a mintából ugyanarra a 30 ismérvre, pontosabban ezek közül a 26 személyi és háztartási ismérvre, amelyekre vonatkozólag az arányos minta kiválasztása során történt ellenőrzés /lásd 2.3 és 2.4 pontok/. Ezeket a becsléseket közvetlenül össze lehetett hasonlítani a megfelelő népszámlálási adatokkal, s így a minta és a súlyozási algoritmus megbízhatóságára ezek a teszt számítások nyujtják a legjobb információt.

A teszt számítások országos eredményeit abszolút számokban a 2. tábla, százalékos megoszlásokra pedig a 3. tábla tartalmazza. A számítások Budapestre és a megyék összesenre is elkészültek, az eredmények hasonlóak az országos adatokhoz. Az adatok azt mutatják, hogy a mintának tulnyomó többségét kitevő azon része, amely az 1980. január 1-én meglevő körzeteket és címeket tartalmazza - a teszt számítások természetesen csak ezekre vonatkozhattak -, a súlyozás figyelembevételével jól reprezentálja a népszámláláskor magánháztartásokban élő népességet. Az abszolút számok esetén a mintából kapott becslések relativ eltérése a népszámlálási adatoktól az esetek felében 1 % alatt van, további 35-40 %-ában 1 és 3 % közötti; a százalékos megoszlásoknál pedig minden esetben legfeljebb tized százalékban van eltérés a mintabeli becslés és a népszámlálási adat között.

Bár a működő ELAR minta tartalmaz még új lakótelepi körzeteket, illetve a régi körzetekben is új építésű lakásokat, aligha valószínű, hogy ezek bekapcsolása a mintába rontja a minta reprezentativitását.

A számítások ugyan csak az "M" részmintára korlátozódtak, de mivel a két részminta nagysága és felépítése teljesen egyforma, az eredmények az egész mintára vonatkoztathatók.

Mint utaltunk rá, a háztartások számától eltekintve a teszt számítások személyi adatokra vonatkoztak, ezek az adatok álltak rendelkezésre a népszámlálási szalagból. Ezekből ugyan a különböző osztályokba, illetve társadalmi rétegekbe tartozó háztartások számára, illetve arányára vonatkozó mintabeli becslések megbízhatóságára nem lehet közvetlenül következtetni, közvetve azonban igen. Teljesen

AZ ÚJ ELAR MINTA M ALMINTÁJA MEGBIZHATÓSÁGÁRA VONATKOZÓ TESZT-SZÁMITÁSOK EREDMÉNYEI  
/körzetenként számítógéppel véletlenszerűen kiválasztott 12 lakás adatainak felszorzása alapján/

Ország összesen  
Abszolút számok

	Férfi					Nő				
	Összesen	0-14	15-39	40-59	60-X	Összesen	0-14	15-39	40-59	60-X
		éves								
1. Minta	5004 861	1181 037	1807 255	1295 269	721 300	5373 783	1102 484	1809 667	1437 779	1023 853
2. Népszámlálás	4987 155	1169 561	1823 214	1252 004	742 376	5390 078	1108 616	1822 367	1410 613	1048 482
3. Eltérés /1-2./	17 706	11 476	-15 959	43 265	-21 076	-16 295	-6 132	-12 700	27 166	-24 629
4. %-os eltérés	0,35	0,98	-0,87	3,45	-2,83	-0,30	-0,55	-0,69	1,92	-2,34

## Aktív keresők

	Férfi		Nő		Összesen	Iparban, építő- iparban	Mezőgazda- ságban	Egyéb ágazatok- ban	Munkás	Szövet- kezeti paraszti	Szellemi	Önálló
	Férfi	Nő	Férfi	Nő								
1. Minta	2802 240	2209 599	5011 839	5011 839	10027 678	2089 147	958 427	1964 265	2818 935	644 847	1419 821	128 236
2. Népszámlálás	2783 703	2175 695	4959 398	2064 486	934 762	2064 486	934 762	1960 150	2800 748	605 290	1415 679	137 681
3. Eltérés /1-2./	18 537	33 904	52 441	24 661	23 665			4 115	18 187	39 557	4 142	-9 445
4. %-os eltérés	0,66	1,55	1,05	1,19	2,53			0,20	0,64	0,53	0,29	-6,86

	Inaktívok		Eltartottak		Lakónépesség	Háztartások száma
	Férfi	Nő	Férfi	Nő		
1. Minta	842 289	1270 343	1360 332	1893 841	10378 644	3723 881
2. Népszámlálás	857 273	1310 992	1346 179	1903 391	10377 233	3719 349
3. Eltérés /1-2./	-14 984	-40 649	14 153	-9 550	1 411	4 532
4. %-os eltérés	-1,74	-3,10	1,05	-0,50	0,01	0,12

Ország összesen  
Százalékos megoszlások

	Férfi						Nő					
	éves			éves			Összesen		Összesen		éves	
	0-14	15-39	40-59	60-X	0-14	15-39	40-59	60-X	0-14	15-39	40-59	60-X
1. Minta	48,22	11,38	17,41	12,48	6,95	51,78	10,62	17,44	13,85	9,86		
2. Népszámlálás	48,06	11,27	17,57	12,06	7,15	51,94	10,68	17,56	13,59	10,10		
3. Eitérés /1-2./	0,16	0,11	-0,16	0,42	-0,20	-0,16	-0,06	-0,12	0,26	-0,24		

	Aktív keresők													
	Férfi		Nő		Iparban, építőiparban		Mezőgazdaságban		Egyéb ágazatokban		Munkás	Szövetkezeti paraszt	Szellemi	Önálló
	Összesen	éves	Összesen	éves	éves	éves	éves	éves	éves	éves	éves	éves	éves	éves
1. Minta	100,0	55,91	44,09	41,69	19,12	39,19	56,24	12,87	28,33	2,56				
2. Népszámlálás	100,0	56,13	43,87	41,63	18,85	39,52	56,47	12,20	28,55	2,78				
3. Eitérés /1-2./	-0,22	0,22	0,22	0,06	0,27	-0,33	-0,23	0,67	-0,22	-0,22				

	Inaktívak						Eltartottak				Átlagos háztartás-nagyság				
	Összesen			éves			Férfi		Nő			Férfi		Nő	
	Összesen	éves	éves	éves	éves	éves	éves	éves	éves	éves		éves	éves	éves	éves
1. Minta	48,29	20,36	8,12	12,24	31,35	13,11	18,24	2,79							
2. Népszámlálás	47,79	20,89	8,26	12,63	31,31	12,97	18,34	2,79							
3. Eitérés /1-2./	0,50	-0,53	-0,14	-0,39	0,04	0,14	-0,10	0							

valószínűtlen ugyanis - és ezt a ténylegesen végrehajtott felvételek eredményei is alátámasztják -, hogy miközben a különböző korcsoportokba, az egyes aktivitási csoportokban, az aktív keresőkön belül a különböző ágazatokba, illetve osztályokba, rétegekbe tartozó személyek számát, illetve arányát megfelelően reprezentálja a minta, az egyes társadalmi rétegek tényleges súlyuktól szignifikánsan eltérő arányban szerepeljenek a felszorozott adatokban.

### 3.6 A MINTA KARBANTARTÁSA, ROTÁCIÓ

A minta karbantartásának célja, hogy biztosítsa, hogy az ELAR minta nemcsak az induláskor, hanem az egész tervezett 10 éves időszak alatt folyamatosan reprezentálja az ország lakásállományát, illetve a magán háztartásokban élő népességet, azaz tükrözze az időközben bekövetkezett változásokat is.

A minta karbantartása kettős feladatot jelent. Egyrészt az egyik évben az "M", a másikban a "H" rész minta körzeteinek teljes cimbejárásával és a cimlista korrekciójával, kiegészítésével biztosítani lehet, hogy a konkrét minták kiválasztásához szolgáló mintavételi keret mindig a mintakörzetek valóságos lakásállományát tükrözze. Ezt a célt segíti elő, hogy a minta települései, illetve körzetei szintjéig bontott teljeskörű építési és bontási statisztikák alapján a Népesedéssziszatizikai főosztály Lakossági Mintavételek osztályán /LAMO/ a minta minden körzetébe évenként beírják az ujonnan épült lakásokat és törlik a lebontott vagy másképpen megszűnt lakásokat. Ezeket a módosításokat kell a kérdőbiztosoknak a helyszíni cimbejárással ellenőrizniük és szükség esetén kiegészíteniük.

A feladat másik része az új lakótelepi építkezések lakásaiból vett megfelelő arányú minta évenkénti hozzákapcsolása az ELAR mintához. Mivel a minta nagysága a megfelelő pénzügyi fedezet hiánya miatt nem növelhető folyamatosan, az induló minta bizonyos számú tartalék körzetet tartalmazott, ezek folyamatos lecserélésével biztosítható az új lakótelepi körzetek bevonása a mintába. Cseréjükig ezek a tartalék körzetek természetesen szerves részei a mintának, azaz minden felvétel ezekben is végrehajtásra kerül, s az adatfeldolgozásban is ugyanugy szerepelnek, mint a többi körzet.

Még egy része van a minta karbantartásának. Az Elnöki Tanács időről-időre módosításokat eszközöl a települések közigazgatási beosztásán: egyes községek városi rangot kapnak, új községek alakulnak, településeket összevonnak, illetve más községekhez, városokhoz csatolnak. Ezeket a változásokat egyrészt érvényesíteni kell a települések rétegcsoportokba történő besorolásánál az egyes rétegcsoportok továbbvezetett teljeskörű lakásszámának megállapításához, másrészt esetenként egyes rétegcsoportok definícióját is módosítani kell, illetve egyeseket megszüntetni vagy újakat alkotni, annak érdekében, hogy a községek és városok különböző rétegcsoportokba kerüljenek.

A rotáció a konkrét felvételekben szereplő címek teljes vagy részleges cseréjét jelenti. Az "M" részmintán 1983-ban és 1984-ben végrehajtott felvételek cimanyaga változatlan volt, ezért 1986-ra az "M" körzetekben teljes címcsereére került sor. A továbbiakban az "M" körzetekben két évenként a címek fele kerül kicserélésre, azaz egy-egy háztartás három évig marad a mintában, de ebből csak két évben szolgáltat adatokat. A "H" körzetekben két évenként a háztartások harmad-része kerül rotálásra, mivel a háztartási költségvetési felvételnél a nagyobb arányú meghiusulás miatt amugyis van egy természetes cserélődése a háztartásoknak.

## II. AZ 1984. ÉVI MIKROCENZUS MINTÁJA

### 1. A MINTAVÉTEL ALAPELVEI

Az 1984. október 1-i eszmei időponttal végrehajtott mikrocenzus - éppugy, mint a korábbi, 1963., 1968. és 1973. évi mikrocenzusok - a népesség 2 %-át volt hivatva felölelni. A korábbi mikrocenzusokhoz hasonlóan most is a szigorú arányosság volt a legfontosabb követelmény a mintával szemben. Arányosságon az értendő, hogy minden területi egységből, rétegből, település nagyságcsoportból egy-egyesen az ott élő népesség 2 %-a kerüljön be a mintába. A mikrocenzusoknál azonban a kiválasztás alapegysége nem a személy, hanem a lakás, így a mikrocenzus mintáját úgy kellett kialakítani, hogy az a folyamatosan továbbvezetett lakásszámokat figyelembe véve rétegenként, kisebb egységenként a lakások 2 %-át ölelje fel. Amennyiben a teljeskörű építési és bontási statisztikákon alapuló továbbvezetett lakásszámok a valóságos lakásállományt tükrözik, a lakásokra vonatkozó 2 %-os minta egyuttal nagy pontossággal a népesség tekintetében is ugyanilyen arányu mintát eredményez.

Mint az I. részben részletesen ismertettük, az új ELAR minta kiválasztásához alapul szolgáló 15 %-os arányos minta minden tekintetben arányos lakásmintának tekinthető - legalábbis az 1982. évi lakásállománynak megfelelően -, ha körzetenként azonos számú lakás kerül kiválasztásra. Így kézenfekvő volt, hogy az 1984. évi mikrocenzus mintájához is ez szolgáljon alapul. Másrészt a felvétel szervezése és minősége szempontjából az is kézenfekvő követelmény volt, hogy a mikrocenzus minél nagyobb arányban az ELAR felvételeknél közreműködő jól képzett és begyakorlott összeírók hajtsák végre. Mint az I. rész 3.2 pontjában szereplő táblázatból látható, a 15 %-os arányos minta közel 300 olyan települést tartalmaz, amelyek nem szerepelnek a működő ELAR mintában. Így a teljes arányos minta felhasználása a mikrocenzus céljára az elsődleges mintavételi egységek nagyobb száma miatt egyrészt ugyan növelte volna az adatok megbízhatóságát, másrészt viszont a felvételbe bevonandó egyszeri alkalomra megbízott összeírók nagy

száma miatt bizonyosan negatívan hatott volna az adatok minőségére. Ennek és a szervezési szempontoknak a figyelembe vételével így olyan döntés született, hogy a mikrocenzus céljára a 15 %-os arányos mintát az arányosság szigorú betartása mellett a településeket tekintve jelentős mértékben tömöriteni kell oly módon, hogy a működő ELAR minta lehetőleg teljesen benne legyen a mikrocenzus mintájában. A működő ELAR mintát természetesen mindenképpen bővíteni kellett, hiszen az nem arányos minta. A tömörités célja éppen az volt, hogy minél kisebb arányban kelljen a működő ELAR mintán kívüli településeket bevonni a mikrocenzus mintájába, s ezzel minél kevesebb egyszeri alkalomra megbízott összeíróra legyen szükség a mikrocenzus végrehajtásához.

A 15 %-os arányos minta tartalmaz minden 10 ezernél népesebb települést, a működő ELAR minta viszont csak a 25 ezernél népesebb települések mindegyikét. A mikrocenzus mintája tekintetében az látszott célszerűnek, ha az kiterjed a 15 ezernél népesebb települések mindegyikére, azaz nemcsak a 25 ezernél nagyobb városok, hanem a 15-25 ezres települések is önreprezentálók lesznek.

Bár a mikrocenzus mintája alapvetően az 1982-ben kiválasztott 15 %-os arányos mintán alapszik, az egyenlőtlen lakásépítési és bontási arányok miatt az arányosságot újra ellenőrizni kellett a rétegenként 1984. október 1-ig továbbvezetett lakásállomány alapján. Ehhez tartozott, hogy az 1983-ban és 1984. I-III. negyedévben épített /illetve várhatóan átadásra kerülő/ lakótelepi lakások reprezentálására ezek 2 %-át tartalmazó új lakótelepi körzeteket kellett kiválasztani a mikrocenzus mintájába.

Lényeges elvi eltérés az arányos minta és a mikrocenzus mintája között, hogy míg az előbbi csak a magán háztartások reprezentálására volt hivatott, az utóbbinak szükségszerűen az intézeti háztartásokra /munkásszállók, kollégiumok, szociális otthonok, csecsemő- és gyermekotthonok/ is ki kellett terjednie, hiszen csak így reprezentálhatta az ország teljes népességét.

## 2. TELEPÜLÉSEK ÉS SZÁMLÁLÓKÖRZETEK KIVÁLASZTÁSA

Település kiválasztáson itt alapvetően az arányos minta települései közül történő kiválasztás értendő, magának az arányos minta településeinek a kiválasztási módszerét az I. rész 2.3 pontja tartal-

mazza. Mint fentebb említettük, az arányos mintából való kiválasztás célja a minta koncentrációja volt. Ennek a koncentrációnak természetesen határt szabott egyrészt az arányosság követelménye, másrészt az a törekvés, hogy a működő ELAR minta településeit lehetőleg a mikrocenzus mintája is tartalmazza. Természetesen az arányosság volt az elsődleges követelmény, amit néhány esetben csak úgy lehetett teljesíteni, ha a működő ELAR minta egyes települései kimaradnak a mikrocenzus mintájából. Országosan összesen 19 ilyen település volt, s ugyancsak az 1984. évi lakásállomány arányainak pontos betartása érdekében 10 olyan települést is be kellett választani a mikrocenzus mintájába, amelyek nem szerepeltek az 1982-ben kiválasztott arányos mintában. Végeredményben a mikrocenzus mintája 524 települést tartalmaz, közel 3/4 részben olyanokat /386-ot/, amelyek a működő ELAR mintában is szerepelnek. A koncentráció révén tehát az arányos minta közel 700 településével szemben a mikrocenzus mintája 174 településsel, 25 %-kal kevesebbre terjedt ki.

Mint az I. részben részletesen ismertettük, mind az arányos mintában, mind a működő ELAR mintában szereplő 4-6. nagyságcsoporthú településekből egységesen hat körzet került a mintába. A mikrocenzus mintájának említett tömörítését a 3-10 ezres településeken úgy lehetett elérni, hogy a mintába került ilyen településeken az eredetileg kiválasztott hat körzeten felül még további körzetek - általában még három körzet - került kiválasztásra. Mivel a mikrocenzus mintájához - a 15 %-os arányos mintával ellentétben - a 10-15 ezres lélekszámú településekből is történt kiválasztás, esetenként itt is szükség volt az arányos minta keretébe nem tartozó körzetek bevonására a mintába. A 15 ezernél népesebb települések esetén viszont a mikrocenzus minta körzetei lényegében megegyeznek a 15 %-os arányos minta körzeteivel, csak néhány esetben került sor kisebb korrekciókra, a lakásállomány időközben bekövetkezett változása miatt az arányosság pontos betartása érdekében. Tulajdonképpen ilyen korrekciót, illetve kiegészítést jelentett az 1983-ban és 1984. I-III. negyedévekben épült és átadott lakótelepi lakások megfelelő arányú reprezentálására új lakótelepi körzetek kiválasztása és bevonása a mikrocenzus mintájába.

Mivel a mikrocenzus mintája a 15 %-os arányos mintán alapul, ahol mind a települések, mind a körzetek nagysággal arányos valószínűséggel kerültek kiválasztásra, így itt is körzeteken belül azonos számú lakást kellett kiválasztani ahhoz, hogy mindenkinek azonos esé-

lye legyen a mintába való bekerülésre, azaz a minta valóban arányos legyen. Figyelembe véve a körzetek átlagos lakásszámát az adódott, hogy a körzetek 15 %-ának kiválasztása mellett körzetenként 14 lakást kell a mikrocenzus mintájába bevonni ahhoz, hogy ez a lakások 2 %-ának kiválasztását eredményezze. Végeredményben így a 3000-nél népesebb települések esetén az egyes területi egységekből, település nagyságcsoportokból, azon belül a települések, illetve körzetek jelleg szerinti rétegeiből kiválasztandó körzetek számát úgy kaptuk, hogy a továbbvezetett, illetve 1984. október 1-re előrebecsült lakásszám 2 %-át osztottuk 14-gyel. Az így adódó körzetszám közelítőleg az adott réteg körzeteinek 15 %-a volt, de a körzetek átlagos lakásszámában mutatkozó eltérések következtében esetenként kisebb-nagyobb mértékben eltért a pontos 15 %-tól.

A 3000-nél kisebb lélekszámú települések esetén a körzetek kis száma miatt általában nem volt mód az arányos mintába nem tartozó további körzetek bevonására a mikrocenzus mintájába, ezért ezen községeknél a minta tömörítését csak a körzeteken belüli több lakás kiválasztása révén lehetett elérni. E településeken ezért egy-egy körzetről nem 14, hanem 20 lakás kiválasztására került sor, s így a mikrocenzus mintája a számlálókörzeteknek csak kb. 10 %-ára terjed ki.

Annak érdekében, hogy a különböző okok miatt óhatatlanul bekövetkező meghiusulások ne csökkentsék a minta méretét 2 % alá, illetve, hogy biztosítani lehessen, hogy a minta megyénként város-község bontásban a továbbvezetett férfi és női népesség pontosan 2 %-át tartalmazza, a mintába eleve bizonyos tartalék került kiválasztásra, mégpedig tartalék körzetek formájában. Ez a tartalék a mintába elsődlegesen kiválasztott körzeteknek átlagosan 6,7 %-át tette ki. A tartalék körzetekben a mikrocenzus ugyanolyan módon került végrehajtásra, mint a többi körzetben. A tartalék körzetekkel együtt a mikrocenzus mintája mintegy 5400 számlálókörzetet ölel fel.

### 3. A MIKROCENZUS MINTÁJÁNAK ÖSSZETÉTELE

Az alábbiakban bemutatjuk, milyen a mikrocenzus település, ill. körzet mintájának a települések jellege és nagysága, illetve megyék szerinti összetétele.

4. táblázat

A mikrocenzus mintája település nagyságkategóriák szerint

Település kategória	Települések	Számláló- körzetek
	száma	
<b>Községek</b>		
1 000-nél kevesebb lakossal	142	286
1 000- 1 499 lakossal	47	171
1 500- 2 999 lakossal	85	470
3 000- 4 999 lakossal	58	445
5 000- 9 999 lakossal	60	504
10 000-14 999 lakossal	23	203
15 000-nél több lakossal	10	90
Községek együtt	425	2 169
<b>Városok<sup>a/</sup></b>		
10 000-nél kevesebb lakossal	6	45
10 000-14 999 lakossal	15	114
15 000-24 999 lakossal	29	303
25 000-nél több lakossal	43	1 086
Városok együtt	93	1 548
Megyei városok	5	478
Budapest	1	1 195
Ország összesen	524	5 390

a/ Az 1984-ben érvényes közigazgatási besorolás szerint.

## 5. táblázat

A mikrocentzus mintájának összetétele megyék szerint

Megyék	Működő ELAR mintából		Az arányos minta nem működő részéből		Arányos mintán kívül		Összesen	
	Települések	Számláló körzetek	Települések	Számláló körzetek	Települések	Számláló körzetek	Települések	Számláló körzetek
	száma							
Baranya	23	75	6	26	-	9	29	110
Bács-Kiskun	26	168	7	104	-	43	33	315
Békés	19	121	6	83	-	35	25	239
Borsod-Abaúj-Zemplén	33	137	10	86	2	32	45	255
Csongrád	15	84	3	56	-	16	18	156
Fejér	14	91	7	85	-	27	21	203
Győr-Sopron	15	73	5	51	1	7	21	131
Hajdu-Bihar	20	107	3	37	-	31	23	175
Heves	13	80	6	65	1	23	20	168
Komárom	12	77	4	68	1	13	17	158
Nógrád	12	57	5	36	1	11	18	104
Pest	33	214	19	184	1	91	53	489
Somogy	25	101	5	41	1	16	31	158
Szabolcs-Szatmár	25	131	9	88	1	36	35	255
Szolnok	17	108	7	92	-	36	24	236
Tolna	16	71	4	38	-	13	20	122
Vas	20	70	5	46	-	11	25	127
Veszprém	22	93	9	75	-	7	31	175
Zala	20	82	8	53	1	6	29	141
Megyék összesen	380	1 940	128	1 314	10	463	518	3 717
Debrecen	1	42	-	64	-	1	1	107
Győr	1	26	-	41	-	2	1	69
Miskolc	1	40	-	67	-	2	1	109
Pécs	1	36	-	56	-	-	1	92
Szeged	1	34	-	62	-	5	1	101
Megyei városok együtt	5	178	-	290	-	10	1	478
Budapest	1	376	-	811	-	8	1	1 195
Ország összesen	386	2 494	128	2 415	10	481	524	5 390

#### 4. LAKÁSOK KIVÁLASZTÁSA

A körzetenkénti 14, illetve 20 lakás kiválasztásához az 1980. évi népszámlálás számlálókörzeti gyűjtőivei /un. "F" minta/ szolgáltak alapul; az 1980. január 1. óta épült lakótelepi lakások reprezentálására kiválasztott új lakótelepi körzetek esetén hasonló címjegyzék készült. A körzetek címjegyzékét naprakész állapotba tették, azaz a folyamatos teljeskörű építési és bontási statisztikák alapján kiegészítették a körzetben 1980. január 1. óta épült lakások címeivel, illetve törölték az időközben lebontott lakásokat. Ezután véletlen kezdőszám és körzetre meghatározott váltószám alapján szisztematikus kiválasztással került sor a 20, illetve 14. mintába kerülő lakás kijelölésére. A működő ELAR minta körzeteiben az 1983., illetve 1984. évi ELAR felvételekben közreműködő háztartások címei - a válaszadók túlzott megterhelésének elkerülése érdekében - nem kerültek be a mikrocenzus mintájába, azaz ha ilyen címre esett a lépés, akkor a címjegyzékben előtte vagy utána szereplő cím került kiválasztásra. A kiválasztás során ellenőrizték és ha szükséges volt, korrekcióval /címcserevel/ biztosították, hogy a minta az 1980. január 1. óta un. szórványban - nem telepszerűen - épített lakásoknak is a 2 %-át tartalmazza.

Végeredményben a mikrocenzus 81 ezer lakás és a benne lakó népesség összeírására terjedt ki.

#### 5. INTÉZETI HÁZTARTÁSOK KIVÁLASZTÁSA

A mikrocenzus nemcsak a magán, hanem az intézeti háztartások népességére is kiterjedt. Az intézeti háztartásokból külön mintát kellett kialakítani a mikrocenzus céljára, hiszen a 15 %-os arányos minta - és következésképpen az ELAR minta is - csak a magán háztartásokban élő népességet reprezentálja. A mikrocenzus intézeti mintája nem terjed ki a kórházakra, üdülőkre, szállodákra, mivel ezekben csak elvétve fordulhat elő állandó vagy ideiglenes bejelentővel ott lakó személy.

A többi intézetnél /csecsemőotthonok, diákotthonok, kollégiumok, munkásszállások, nővérszállások, szociális otthonok és otthonházak/ típus és azon belül nagyságkategóriánként rétegezve történt

a kiválasztás. A nagyságkategóriák kialakítása az intézet befogadó-képessége /férőhelye/ alapján történt, tipusonként némileg eltérő nagyságkategóriákat alkalmazva. A kiválasztás a legnagyobb nagyság-kategória kivételével kétlépcsős volt, első lépcsőben került sor az intézet kiválasztására, majd ezeken belül az összeíró a központilag előírt kezdőszám és váltószám alapján választotta ki az intézetben lakók névsorából a mintába kerülő személyeket. Minden intézet típusnál a legnagyobb intézetek mind szerepeltek a mintában, ezeknél csak intézeten belül történt kiválasztás. A mintavételi terv, amely nagyságkategóriánként megadta a kiválasztási arányt az intézetekből és intézeten belül, tipusonként némileg eltérő volt, mivel a különböző típusu intézetek összetétele a férőhely nagysága szerint erősen különbözik. A két kiválasztási arány szorzata minden típusnál és nagyságkategóriában természetesen kiadta a 2 %-ot.

Az intézeti háztartások mintája is tartalmaz tartalékot, mégpedig 10 %-ot. A tartalék a kiválasztott intézetekből további személyek mintába vevésével adódott.

## 6. A MINTA ELLENŐRZÉSE ÉS KORREKCIÓJA

A mikrocenzus mintájából közvetlenül az adatok rögzítése után készült egy nyers feldolgozás annak ellenőrzésére, mennyire egyeznek meg az eredeti - tartalék nélküli - mintából kapott személyszámok nemek és megyék szerint, azon belül város-község bontásban a születési, halálozási és belső vándorlási statisztikákra, illetve a népességnyl-vántartásra támaszkodó továbbvezetett népesség adatokkal.

Ahol eltérés mutatkozott, ott a tartalékként felvett anyag felhasználásával korrigáltuk a mintát oly módon, hogy mindkét nemre, megyéken belül külön a városokra és községekre a minta személyszámai pontosan a továbbvezetett előzetes népességszámok 2 %-át adják, azaz az egységes 50-es szorzószámot alkalmazva a minta személyszám tekintetében a jelzett bontásban pontosan tükrözze a továbbvezetett lakó-népességi adatokat.

A korrekció során tulnyomórészt az alábbi eljárások kerültek alkalmazásra:

- Teljes tartalékkörzetek bevonása a mintába
- Teljes körzetek elhagyása a mintából

- Tartalékkörzetek egyes háztartásainak bevonása a mintába
- A mintában lévő egyes háztartások elhagyása

Megjegyzendő, hogy a korrekció során sohasem került sor háztartások szétszakítására. Nyers feldolgozás és ellenőrzés nem, kor és családi állapot szerint is történt, és a korrekciónál lehetőleg figyelemmel voltunk a kormegoszlásnál mutatkozó eltérésekre a minta és a továbbvezetett adatok között.