

# STATISZTIKAI SZEMLE

A KÖZPONTI  
STATISZTIKAI HIVATAL  
FOLYÓIRATA

SZERKESZTŐBIZOTTSÁG:

DR. BELYÓ PÁL, ÉLTETŐ ÖDÖN, DR. HARCSA ISTVÁN,  
DR. HUNYADI LÁSZLÓ (főszerkesztő), DR. HÜTTL ANTÓNIA, DR. KÖRÖSI GÁBOR,  
DR. MÁTYÁS LÁSZLÓ, DR. MELLÁR TAMÁS (a Szerkesztőbizottság elnöke), NYITRAI FERENCNÉ DR.,  
OROS IVÁN, DR. RAPPAI GÁBOR, DR. SIPOS BÉLA, DR. SZILÁGYI GYÖRGY,  
TÓTH ISTVÁN GYÖRGY, DR. VITA LÁSZLÓ, DR. VUKOVICH GABRIELLA

80. ÉVFOLYAM 4. SZÁM

2002. ÁPRILIS

## E SZÁM SZERZŐI:

*Dr. Fóti János*, a KSH ny. fősztályvezető-helyettese; *Herman Sándor* kandidátus, a Pécsi Tudományegyetem tanszékvezető docense; *Dr. Hunyadi László* kandidátus, a Budapesti Közgazdaságtudományi és Államigazgatási Egyetem egyetemi tanára, a *Statisztikai Szemle* főszerkesztője; *Dr. Lakatos Miklós*, a KSH fősztályvezető-helyettese; *Németh Renáta*, a Johan Béla Országos Epidemiológiai Központ felmérésstatistikusa; *Dr. Révész Tamás*, a Gazdasági Minisztérium fősztályvezető-helyettese; *Rudas Tamás* kandidátus, az Eötvös Loránd Tudományegyetem Szociológiai Intézetének igazgatója; *Dr. Varga Attila*, a Pécsi Tudományegyetem docense; *Vargha András* kandidátus, az Eötvös Loránd Tudományegyetem docense.

\*

*Balogh András* kandidátus, a KSH főtanácsosa; *Hajnal Béla* kandidátus, a KSH Szabolcs-Szatmár-Bereg Megyei Igazgatóság igazgatója; *Hejna Ferencné*, a KSH főtanácsosa; *Szász Kálmán* kandidátus, a KSH Könyvtár és Dokumentációs Szolgálat tudományos kutatója; *Tűz Lászlóné*, a KSH ny. osztályvezetője.

---

ISSN 0039 0690

---

Megjelenik havonta egyszer  
Főszerkesztő: dr. Hunyadi László  
Osztályvezető: Dobokayné Szabó Orsolya  
Kiadja: a Központi Statisztikai Hivatal  
A kiadásért felel: dr. Mellár Tamás  
3501 – Akadémiai Nyomda  
Martonvásár, 2001  
Felelős vezető: Reisenleitner Lajos

---

Szerkesztők: dr. Domokos Attila, Polyák Andrea, Szűcsné Bruckner Mariann, Visi Lakatos Mária  
Tördelőszerkesztők: Bálinthné Bartha Éva, Simonné Káli Ágnes

---

Szerkesztőség: Budapest II., Keleti Károly utca 5–7. Postacím: Budapest, 1525. Postafiók 51.  
Telefon: 487-4341, 487-4343 Telefax: 487-4344  
Internet: [www.ksh.hu/statszml](http://www.ksh.hu/statszml)  
E-mail: [statszemle@ksh.gov.hu](mailto:statszemle@ksh.gov.hu)

Kiadóhivatal: Központi Statisztikai Hivatal, Budapest II., Keleti Károly utca 5–7.  
Postacím: Postafiók 51. Budapest, 1525. Telefon: 345-6000

Előfizetésben terjeszti a Magyar Posta Rt. Előfizethető bármely hírlapkézbetítő postahivatalnál és a Üzleti és Logisztikai Központ Hírlapelőfizetési Irodájánál (Budapest VIII., Orczy tér 1., Telefax: 303-3440) közvetlenül vagy postautalványon, valamint átutalással Postabank Rt. 219-98636, 021-42795 pénzforgalmi jelzőszámra.  
Előfizetési díj: fél évre 3000 Ft, egy évre 5400 Ft

Beszerezhető a KSH Könyvesboltban. Budapest II., Keleti Károly u. 10. Telefon: 212-4348

## TARTALOM

### MÓDSZERTANI TANULMÁNYOK

Mintavétel a Leslie Kish-kulcs alkalmazásával. – <i>Németh Renáta – Rudas Tamás</i> .....	309
Független minták összehasonlítása új rangsorolási eljárásokkal. – <i>Vargha András</i> .....	328
Térökonometria. – <i>Varga Attila</i> .....	354

### STATISZTIKAI ELEMZÉSEK

A háztartási jövedelmek és adóterhek rétegmegoszlása. – <i>Dr. Révész Tamás</i> .....	371
A munkaerő-piaci helyzet a 2001. évi népszámlálás tükrében. – <i>Dr. Fóti János – Dr. Lakatos Miklós</i> .....	394

### SZEMLE

Évnyitó értekezet a Központi Statisztikai Hivatalban. – <i>H. L.</i> .....	418
A statisztika hete Bécsben. – <i>Herman Sándor</i> .....	421

### STATISZTIKAI HÍRADÓ

Személyi hírek .....	423
Szervezeti hírek – Közlemények .....	424

### STATISZTIKAI IRODALMI FIGYELŐ

#### Külföldi statisztikai irodalom

Blackwell, L.: A nők munkavégzése az Egyesült Királyságban és a foglalkozások újraosztályozása. ( <i>Tűű Lászlóné</i> ).....	427
Omar, E.: A statisztika iránti növekvő igény az emberi jogok területén. ( <i>Szász Kálmán</i> ) .....	429
Ayhan, H. Öztas: A nemek asszimmetriájának csökkentése a mezőgazdasági összeírásokban. ( <i>Balogh András</i> ).....	430

Söndermann, M.: A kulturális ágazatok és a foglalkoztatás Európában. (Hajnal Béla).....	431
Rothstein, D. S.: A fiatalok foglalkoztatottsága az Egyesült Államokban. (Hejna Ferencné).....	434
Külföldi folyóiratszemle .....	436

*A Statisztikai Szemlében megjelenő tanulmányok  
kutatói véleményeket tükröznek, amelyek nem esnek szükségképp egybe  
a KSH vagy a szerzők által képviselt intézmények hivatalos álláspontjával.*

*Utánnnyomás csak a forrás megjelölésével!*

### MINTAVÉTEL A LESLIE KISH-KULCS ALKALMAZÁSÁVAL\*

NÉMETH RENÁTA – RUDAS TAMÁS

A megkérdendő személy háztartáson belüli kiválasztásának problémája gyakran merül fel a többlépcsős mintavételi eljárás utolsó lépcsőjében; például telefonos felmérésekben, miután a telefonszámok véletlen kiválasztásával háztartási mintához jutottunk. A Kish-kulcs a háztartáson belüli választásra ad eljárást. Eredményeink szerint, bár a kulcs a háztartáson belül véletlen mintát biztosít, az általánosan elterjedt nézettel szemben nem alkalmas a sokaságban jellemző nemek és korcsoportok szerinti arány beállítására. A tévedés feltehetőleg abból származik, hogy amikor Kish az 1950-es évek Amerikájában az eljárást kidolgozta, az ország aktuális háztartásszerkezete miatt ez a két cél egyszerre volt megvalósítható. Számításaink szerint a mai magyarországi helyzetben ez nincs így. A Kish-kulcs olyan módosítását javasoljuk, mely alkalmasabb reprezentatív minták kiválasztására.

TÁRGYSZÓ: Felvételtervezés. Mintavétel. Leslie Kish-kulcs.

A Leslie Kish-kulcsot vagy módosított változatait gyakran alkalmazzák hazai és külföldi felmérésekben.<sup>1</sup> A kulcs ugyanúgy a háztartáson belüli véletlen kiválasztásra szolgál, mint például a „legutóbbi születésnap” módszere. Alkalmazásakor a mintavétel során először a háztartásokat választjuk ki, majd a mintába került háztartásokon belül jelöljük ki a kérdezendő személyt. Az 1. részben tárgyaljuk azokat a tényezőket, amelyek ennek a kétlépcsős mintavételi módszernek az alkalmazását indokolhatják.

\* A tanulmány alapjául *Németh Renáta* egyetemi szakdolgozata szolgált, melyet az ELTE Szociológiai Intézetének szociológus szakán készített, *Rudas Tamás* témavezetésével.

<sup>1</sup> Néhány példa: 1. *A közigazgatás megítélése* (1999). MONITOR Társadalomkutató Intézet és Módszertani Központ. <http://www.monitor-tki.hu>. 2. *Csaba TV nézettségi mutatók* (1998). Jelenkutató Intézet. <http://www.bekes.hungary.net/csabav/nezetseg.htm>. 3. *Felmérés Veresegyházán és környékén* (1998). (Táblaképek az egészségről – A veresegyházi példa) (2001). 4. *ISSP* (1994). Család II., TÁRKI. <http://www.tarki.hu>. 5. Az Egyesült Államokban: telefonos politikai közvéleménykutatások 1998-ban és 1992-ben, a felméréseket vezette a CBS és a New York Times. (*Voss-Gelman-King*, 1995). 6. National fertility and family survey Európa több országában, 1990-es évek. Kivitelező: Population Activities Unit of the United Nations Economic Commission for Europe. <http://www.cpc.unc.edu/pubs/paa/papers/1995/klijzing.html>. 7. International Social Justice Project, 1991, 1996. 12 európai ország, köztük Magyarország bevonásával. <http://www.isjp.de/>. 8. *SCSU Survey*, Minnesota State Lottery (1997). <http://www.lottery.state.mn.us/gambling/stcloud.html>. 9. *Los Medios y Mercados de Latinoamérica* (1997), több latin-amerikai országban. <http://www.zonalatina.com/Zldata25.htm>. 10. *Northern Ireland social attitudes survey* (1989–1996). Social and Community Planning Research <http://cain.ulst.ac.uk/othelem/research/nisas/robinson.htm>. 11. *Scottish health survey* (1995). Social and Community Planning Research <http://www.official-documents.co.uk/document/scottish/shealth/shch1.htm#1.2.1>. 12. *Oregon health behavior surveys* (1998). State of Oregon, Health Division <http://pub.das.state.or.us/purchasing/contracts/8042.html>. 13. *Australian national social science survey*, 1984. The Australian National University <http://ssda.anu.edu.au/SSDA/CODEBOOKS/FR84/description.html>.

A tanulmány középpontjában a Kish-kulcs használatával kapott minták nem és kor szerinti reprezentativitása áll. A Kish-kulcs szokásos használatát a 2. részben ismertetjük. Tapasztalatok szerint a minták nem megfelelően illeszkednek a sokasághoz, és az eltérések szisztematikus hasonlóságot mutatnak.<sup>2</sup> A 3.1. részben kifejtett hipotézisünk szerint ezek az eltérések levezethetők magából a mintavételi eljárásból, s magyarázatukkor nincs szükség a megvalósítás során adódó problémák (például szisztematikus válaszmegtagadás) feltételezésére. A hipotézis alátámasztásaként egy magyarországi felmérésből származó mintát veszünk példának a 3.2. részben. A példa esetlegességén túlmutatva a 3.3. részben a hipotézis bizonyítására is alkalmas számítást közlünk. A bizonyítás során az elméletileg várható minta összetételét határozzuk meg. Ugyanitt rámutatunk arra, hogy nem a kulcs okozza a hibát; ugyanez a probléma lép fel minden, háztartásmintára épülő módszer esetén. A kulcs esete azért is szerencsés, mert használatakor ellenőrizhető a minta nem és kor szerinti összetétele. Ezt kihasználva hajtjuk végre a kulcs módosítását a 4. részben. Eredményünk szerint a tökéletes reprezentativitás még elméletileg sem érhető el, de a minta illeszkedése javítható.

### 1. A HÁZTARTÁSMINTA HASZNÁLATÁNAK OKAI

A felvétel kivitelezésekor a kutatási terv létrehozása az első lépés. A kutatási tervben a felvétel tárgyát képező egyedek összességét, a célsokaságot, és a megismerni kívánt célsokasági jellemzőket határozzuk meg. Ezen kívül a terv rögzíti a célsokaság gyakorlati elérésére alkalmas keretet. A keret köztes állomás a célsokaság és a minta között, azaz a mintát a keretből választjuk.

A mintavétel módjának meghatározása ezen paramétereknek megfelelően történik, figyelembe véve a felvétel kivitelezésének költségvonzatát. A megfelelő eljárás kiválasztása minden felvétel tervezésekor egyedi döntést igényel.

A háztartáson belüli kiválasztás problémája olyan kétlépcsős mintavételi eljárás alkalmazásakor merül fel, amikor az első lépcsőben a mintába kerülő háztartásokat választjuk ki a keretből, majd a háztartás (felölt) tagjai közül jelöljük ki a kérdezendő személyt a második lépcsőben.<sup>3</sup>

Miért alkalmazunk két lépcsőt, miért nem választjuk ki a keretből egy lépésben, közvetlenül a kérdezetteket? Ennek gyakori oka, hogy nem áll rendelkezésre megfelelő, keretként használható lista a célsokaságról. A választói névjegyzékek adatai például több országban sok hibás adatot tartalmaznak a populáció mobilitása vagy a nyilvántartásba nem vétel miatt. Ugyanakkor, bár nem megfelelő a minőségük a személyek kiválasztásához, de eredményesen használhatók címlistaként, mert hibáik a személyek és lakcímek

<sup>2</sup> A magyar statisztikai irodalom a sokaság mellett a populáció kifejezést is használja.

<sup>3</sup> A háztartás mint mintavételi egység gyakran fordul elő a felmérésekben. Leggyakrabban az együtt élők olyan közösségét értik alatta, akik funkcionálisan is együttműködnek. Egyik ilyen megfogalmazás szerint a részben közös jövedelemből élő, közös konyhát használók tartoznak egy háztartásba. Hasonló mintavételi egység lehet a család. A család általában a rokoni kötélekkel is összekapcsolt együtt élőket jelentheti, de például a Központi Statisztikai Hivatal meghatározása szerint a család házastársi vagy élettársi kapcsolatban álló, gyermeces vagy gyermek nélküli személyeket jelöl, illetve egyedül álló szülőt nőlen/hajadon gyermekkel. A család és a háztartás itt tehát nem teljesen átfedő kategóriák, az egy fedél alatt élő háztartás több generációja külön családokat alkot. Ugyanígy létezik a szakirodalomban a „nem családháztartás” kifejezés, ami egymással családi kapcsolatban nem álló együtt élőket jelent. Egy harmadik, a háztartás és a család fogalmát megint csaknem átfedő fogalom az együtt, egy lakásban élő személyek összessége. Problémafelvetésünk szempontjából nem okoz nehézséget a három fogalom megkülönböztetése, hisz bárhol is határozza meg a felméréstervező a mintavételi egységet, ugyanaz a probléma jelentkezik mindhárom esetben.

összekapcsolásában jelentkeznek, de magukat a lakcímekeket megfelelő minőségben tartják nyilván.<sup>4</sup> Egyes felmérésekben ezért a nyilvántartási listát a lakások vagy háztartások kiválasztásához használják, míg a kérdezett kiválasztását a háztartásból újabb véletlen mintavétellel oldják meg.<sup>5</sup>

A megfelelő nyilvántartások hiányának kezelésére fejlesztettek ki egy másik többlépcsős módszert, a területi mintavételt (area sampling) (*Kish*; 1965). Alkalmazásának feltétele, hogy a célsokaság egy adott földrajzi egység lakosságaként legyen meghatározható. A területet blokkokra osztják (például kerületek), ezeket kisebb egységekre (utcák). A területi módszer lényege, hogy földrajzilag definiált mintavételi egységeken működik, így jó területi diszperzitást biztosít. Végső lépcsőben (például véletlen sétával) itt is háztartási mintához jutunk, és a kérdezendőt újabb mintavétellel választjuk.

A telefonos felvételek során is felmerül a háztartások véletlen kiválasztása után a kérdezett kiválasztásának problémája. Ilyenkor a háztartások kiválasztásának lépcsőfoka a mintába kerülő telefonszámok kiválasztásának feleltethető meg.

Tehát a második lépcsőben a háztartás tagjai közül kell kiválasztanunk a kérdezendő személy(eke)t. Nem merülne fel a kérdezendő kiválasztásának problémája, ha a felvétel kutatási céljának megfelelően ő a háztartás meghatározott tagja, például a családfő lenne. Ugyanígy nem okozna nehézséget a választás, ha a mintavételi egység maga a háztartás lenne, ekkor bármely háztartástag egyaránt adekvát információkat szolgáltatathatna.

Általában a háztartások több tagja tartozik a kerethez. A kiválasztás ekkor sem okoz nehézséget, ha az alábbi feltételek bármelyike teljesül.

a) Bármelyik személytől beszerezhetők az összes többi háztartásbelire vonatkozó szükséges információk.

b) A háztartás minden tagját bevonva a felmérésbe sem csökken jelentősen a becslések megbízhatósága az azonos mintanagyság mellett, de egyszerű véletlen mintavétellel kapott mintából számolt becslések megbízhatóságához képest.<sup>6</sup> Ilyenkor a felmérés kérdéseire adott válaszoknak elhanyagolható a háztartásokon belüli korrelációja. Ellenkező esetben, vagyis amikor egyfajta homogenitás tapasztalható a háztartásokon belül, megbízhatóbb becsléseket kapunk, ha (azonos mintanagyság mellett) háztartásonként csak egyetlen személyt választunk.<sup>7</sup>

<sup>4</sup> Hazánkban a Központi Nyilvántartó és Választási Hivatal listája alkalmazható a személyi minta kiválasztására. A TÁRKI tapasztalatai szerint (<http://www.tarki.hu/kozvelemeny-h/adatfelv/metodol.html>) a Hivatal listájának alkalmazásakor a tervezett minta körülbelül 55 százalékának sikeres lekérdezésére lehet számítani. Ebben a 45 százalékos kiesési arányban a címlista hibái mellett persze más okok, így a válaszmegtágadás is közrejátszik. Úgy találták, hogy a kiesés szisztematikus torzítást okoz, a megmaradó minta demográfiai összetétele eltér a kiinduló mintától. Másik tapasztalatra is hivatkozhatunk: az Országos Lakossági Egészségfelmérés 2000 szintén a Hivatal címlistájára támaszkodott. (OLEF2000. A kutatást az Egészségfejlesztési Kutatóintézet vezette, a terepmunka kivitelezője a Magyar Gallup Intézet volt. Az adatfelvétel időpontja 2000. november-december. A Központi Népegységyilvántartó és Választási Hivatal címlistájából készült országos minta. Módszere: kérdőív kérdőbiztosokkal. Esetszám: 5503.) A sikertelen interjúk aránya 21 százalékos, ebből 7 százalék adódott a címlista valamely hibája miatt (például a kérdezendő elköltözött, vagy meghalt). A kiinduló minta ezen 7 százaléka szignifikánsan eltér a teljes mintától a személyek életkorát tekintve (a kiesők főként fiatalok), illetve a személyek lakhelye szerint (több köztük a budapesti lakcímmel szereplő).

<sup>5</sup> Például Nagy-Britanniában: 1. lábjegyzet 11-es felmérése, illetve lásd a Market research methods forrást.

<sup>6</sup> A becslés érvényessége (validity) arra utal, vajon a többszöri mintavétel után kapott becsléssorozat mennyire mozog a valódi sokasági paraméter körül. Ha az ingadozás centruma egy másik érték, az eljárás torzított. Az érvényesség mértéke a torzítás segítségével ragadható meg: a torzítás csökkenésével nő a becslés érvényessége. A becslés megítélésének másik szempontja a megbízhatóság (reliability). A becslés megbízhatósága attól függ, mennyire hasonló becsléseket kapnánk a mintavétel többszöri megismétlésével. A megbízhatóság a becslés varianciájának segítségével mérhető: minél kisebb a varianciája, annál megbízhatóbb a becslés. Fontos megjegyezni, hogy mindkét párhuzam (torzítás-érvényesség, illetve variancia-megbízhatóság) csak akkor érvényes, ha feltesszük, hogy nem lép fel nem-mintavételi hiba (*Levy-Lemeshow*; 1999).

<sup>7</sup> Ezen homogenitás mértéke az osztályon belüli korreláció (intra-class correlation) együtthatója, a  $\rho$ . A csoportos mintavétellel kapott mintából számolt becslés varianciája növekszik, ha a  $\rho$  nő, illetve ha nagyobb a csoportok átlagos nagysága. A variancia növekedése kevésbé megbízható becsléseket eredményez (*Kish*; 1965).

Ezek a feltételek kevés felmérés esetén teljesülnek. Tehát egy személyt kell háztartásonként kiválasztani. Rögzítsük továbbá, hogy feleslegesen nem keresünk fel háztartásokat: minden mintaháztartásból kiválasztunk egy kérdezendőt. Ennek az előfeltételnek a költségkímélés az ésszerű indoka. Vegyük még fel feltevéseink közé, hogy a háztartások egyenlő valószínűséggel kerültek a mintába.

A kérdezés gyakorlatában a kontrollálatlan kiválasztás lenne a legegyszerűbb módszer, amikor is azt a személyt választanánk, aki kinyitja az ajtót, vagy – telefonos felmérésben – aki felveszi a telefont. Ám ez a kiválasztás erős torzítást eredményezne a mintában, mivel azokat kérdeznénk, akik könnyen elérhetők, illetve akik inkább hajlandók együttműködni.

## 2. A KISH-KULCS

Leslie Kish<sup>8</sup> szándéka szerint a kulcs egy világos és előre rögzített eljárást ad a kérdező kiválasztására (Kish; 1949, 1965). Lényege, hogy segítségével a véletlen kiválasztást jól imitáló módon választhatjuk ki a háztartás egyik tagját. Másik előnye, hogy használatának megfelelő dokumentációja birtokában a választás utólag is ellenőrizhető.

A Kish-módszert alkalmazó felmérésben minden kérdőívhez egy fedőlapon csatolnak. A fedőlapon a háztartástagok felsorolására alkalmas tábla (példa az 1. tábla) és egy kiválasztási tábla (példa a 2. tábla) van. Az 1. tábla egy, a kérdező által a terepmunka során már kitöltött táblát illusztrál.

1. tábla

Kiválasztási adminisztráció a terepmunka során

Rokonsági viszony a családfőhöz	Nem	Kor	Sorszám	Kiválasztás
Családfő	F	40	2	
Feleség	N		5	
Családfő apja	F		1	
fia	F		3	
lánya	N		6	
A feleség nagynénje	N	44	4	✓

Forrás. Itt és a 2–5. táblánál: Kish; 1965.

A háztartás tagjai közül elsőként a férfiakat állítjuk sorba, életkor szerint (az idősektől a fiatalabbak felé), majd utánuk a nőket soroljuk fel, szintén életkoruk alapján. Az interjú első néhány kérdése az ehhez szükséges információkat gyűjti össze. Célszerű minél kevesebb kérdésből megállapítani a sorrendet, ezt segíti a rokonsági viszony megállapítása. Az 1. táblabeli példában csupán a feleség és a feleség nagynénje életkorának ismeretében

<sup>8</sup> Leslie Kish (1910–2000) magyar származású amerikai statisztikus volt. A világ elsőszámú mintavételi szakértőjeként jegyezték. 1965-ös könyve, a „Survey Sampling”, ma is használt forrás szerte a világon. Véletlen mintavételi technikájának megalapozottsága először az 1948-as amerikai elnökválasztáskor nyert bizonyítást. (Kevesebb, mint 1000 háztartást tartalmazó országos mintája Dewey és Truman szoros eredményét jósolta, kis fölénnyel Truman javára, míg a sajtó és a szokott felmérések Dewey megsemmisítő győzelmét várták.) Számos eredményéből a legfontosabbak közé tartozik a választásmegtagadási arány figyelembevételének szükségessége a valódi véletlen minta eléréséhez; vagy a tervezéshatás-együttható megalkotása, amely alkalmas a felmérések hatékonyságának mérésére. Kish az elsők között javasolta a guruló minták évenkénti kivitelezését.



végrehajtható volt a sorrendezés. A felsorolt háztartástagok közül a kiválasztandó személyt a Kish-tábla jelöli ki. Példánk esetén 6 felsorolt van, a 2. tábla szerint ekkor a sorban a 4. személy a kérdezendő. (Ha a kiválasztott nem található otthon, a kérdező újra felkeresi a háztartást.)

2. tábla

*A 8 Kish-tábla egyike – D kiválasztási tábla*

Felnőttek száma a háztartásban	A kiválasztott sorszáma
1	1
2	2
3	2
4	3
5	4
6 vagy több	4

A D kiválasztási tábla csak egyike a nyolc hasonló táblának. (Lásd a 2. táblát.) Ezeket a táblákat a kérdőívekhez csatolják, véletlenszerűen párosítva őket. A nyolc tábla nem egyenlő arányban oszlik meg a kérdőívek között, például a C jelzetű minden 6. kérdőíven található meg, míg az E csak minden 12.-en. Az elsődleges cél a háztartásokon belül egyenlő kiválasztási valószínűséget rendelni minden személyhez (anélkül hogy túl sok táblát kellene létrehozni). A 3. táblát elemezve kitűnik, hogy ez csaknem minden háztartás esetén megvalósul. A 3. tábla alapján határozhatók meg a 4. táblában található háztartáson belüli kiválasztási valószínűségek. Például a 3 felnőtt tagot számláló háztartások sorban 1. tagját  $1/6+1/12+1/12=1/3$  valószínűséggel, a 2. tagját  $1/6+1/6=1/3$  valószínűséggel, a 3. tagját  $1/12+1/12+1/6=1/3$  valószínűséggel választjuk ki. Az egyenlő esélyű kiválasztás két esetben nem valósul meg. Az egyik a 6-nál több felnőtt tagot számláló háztartások esete. A 7. vagy nagyobb sorszámú tag 0 valószínűséggel kerül a mintába. A másik eltérés az 5 felnőtt tagú háztartásoknál figyelhető meg. Ott az 1. 2. és 4. tag  $1/6$  valószínűséggel kerül kiválasztásra, a 3. és az 5.  $1/4$  valószínűséggel. A várhatóan fiatal, nőnemű 5. személy felülreprezentálása Kish szerint ellensúlyozza a 6 tagú háztartásokból 0 valószínűséggel bekerülő, várhatóan szintén fiatal lányok kiesését.

3. tábla

*A Kish-kulcs kiválasztási szabálya*

A tábla aránya	A tábla jelzete	A kiválasztott sorszáma, ha a felnőttek száma					
		1	2	3	4	5	6 vagy több
1/6	A	1	1	1	1	1	1
1/12	B1	1	1	1	1	2	2
1/12	B2	1	1	1	2	2	2
1/6	C	1	1	2	2	3	3
1/6	D	1	2	2	3	4	4
1/12	E1	1	2	3	3	3	5
1/12	E2	1	2	3	4	5	5
1/6	F	1	2	3	4	5	6

4. tábla

*Háztartáson belüli kiválasztási valószínűségek a Kish-kulcs alkalmazása esetén*

A tag sorszáma a sorban	A háztartástagok száma					
	1	2	3	4	5	6 vagy több
1	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$
2		$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$
3			$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$
4				$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$
5					$\frac{1}{4}$	$\frac{1}{6}$
6						$\frac{1}{6}$
7 vagy nagyobb						0

A Kish-féle módszer elsősorban a véletlen kiválasztást imitálja. A leggyakrabban a telefonos felmérésekben is ezzel az eljárással választják ki a kérdezettet, bár adott a technika a véletlenszám-generálásra, és ezért a kiválasztásra elegendő információ lenne a háztartástagok azonosítása, például keresztnév alapján. Vélhetően azért ragaszkodnak ma, a számítógépes háttértechnika birtokában is, a nem/kor alapján történő felsoroláshoz, mert ez a módszer bizonyos mértékig biztosítja – a szisztematikus mintavételhez hasonlóan – a minta nem- és korösszetételének kiegyenlítetttségét (Kish maga nem említi ezt az előnyt).

### 3. A MINTA REPRESENTATIVITÁSA

A mintavételi módszer megítélésének egyik szempontja a mintából levezethető becslések megbízhatósága. A becslés megbízhatósága (reliability) attól függ, mennyire hasonló becsléseket kapnánk a mintavétel többszöri megismétlésével. A megbízhatóság szempontjából a minta kívánatos tulajdonsága a gyakran emlegetett reprezentativitás. A reprezentativitás szó jelentésének nagyszámú változatával találkozhatunk. A legtágabb értelemben az adatok jóságát, a minta tipikus voltát értik rajta. Szűkebben a sokaság valamely jellemzője szerinti megoszlásának mintabeli reprodukálását jelenti. Ez azért kívánatos, mert ha a kutatás kérdései összefüggésben állnak az adott jellemzővel, akkor a becslés várhatóan megbízhatóbb lesz a jellemző szerinti reprezentativitás esetén.

A Kish-kulcsot vagy módosított változatait használó hazai és külföldi felmérések során a minták jellemzésekor azonos problémák tapasztalhatók (Kennedy; 1993, *Táblaképek az egészségről...*; 2001, az 1. l. ábrájegyzet 4. és 8. felmérése). Gyakori a nők felülreprezentáltsága és az idősek, főként az idős nők vártnál nagyobb mintabeli aránya. A hozzáférhető források az esetek legnagyobb hányadában az eltéréseket megvalósíthatósági problémákkal magyarázzák, azaz bizonyos demográfiai jellemzők szerinti szisztematikus válasziánnyal. Hasonlóan, amikor Leslie Kish az általa létrehozott eljárással kapott minták reprezentativitását értékelte, a férfiak alulreprezentáltságát ritka otthon tartózkodásukkal és nagyobb arányú válaszmegtagadásukkal magyarázta (Kish; 1949).

#### 3.1. A reprezentativitás sérülésének feltételezhető fő oka

A nehezen ellenőrizhető gyakorlati okokon kívül más forrása is lehet a problémának. A kulcs a háztartáson belüli egyenlő esélyű kiválasztást hivatott teljesíteni, és ez nem je-

lenti a mintába kerülési valószínűségek egyezését. A nagyobb létszámú háztartásokból nyilván nehezebb a mintába kerülni, míg az egyszemélyes háztartások lakóját mindenképpen megkérdezzük, ha háztartása már bekerült a háztartásmintába.

Az általunk felkutatott források jó része a minta reprezentativitásának értékelésekor nem veszi figyelembe az eltérő bekerülési esélyből fakadó elvi szintű problémákat. Ha a háztartások nagyságától függ a bekerülési esély, a háztartások nagysága viszont karakteres jellegzetességeket hordoz bizonyos demográfiai jellemzők szerint, akkor maga a mintavétel okozza a szisztematikus demográfiai eltéréseket. Vagyis ekkor akár tökéletesen véletlen háztartásminta, és nem szisztematikus válaszmegtagadás mellett sem kapunk reprezentatív mintát. Szélsőséges példával: akkor is „rossz” mintát kapunk, ha minden célsokaságbeli háztartást bevonunk a mintába (kiküszöbölve a háztartásválasztás mintavételi hibáit), és teljes együtműködést tételezünk fel a kérdezettek részéről.

A kiválasztási valószínűségek kiegyenlítését Kish súlyozással javasolja megoldani. A kulccsal kapott minták súlyozásáról Kish megállapította, hogy nem okoz jelentős változást a becslésekben, ezért a gyakorlatban elhagyták az elemzésből (*Kish*; 1949, 1965). Kish tapasztalata a súlyozás elhagyhatóságáról a korabeli egyesült államokbeli háztartásszerkezetből eredeztethető. Más körülmények között – más országban, más korban – a minta elégtelen reprezentativitása esetleg már magából a mintavételi módszerből levezethető, és nem csak a szisztematikus válaszmegtagadás az előidézője a torzulásoknak.

A háztartásmintán alapuló felmérések mintája és a sokaság közötti nem és kor szerinti eltérés visszavezethető a háztartásstruktúrára, ha a háztartások nagysága összefügg nemek szerinti és korösszetételükkel. Ekkor a háztartáslétszámmal fordítottan arányos kiválasztási valószínűség torzítja a minta e két jellemző szerinti összetételét. Az alábbiakban ezt a feltevést igazoljuk.

A Központi Statisztikai Hivatal 1996-os mikrocenzusa szerint 1996-ban a háztartások több mint negyede (26%) egyetlen lakóval rendelkezett. A népesség öregedési folyamata megváltoztatta a háztartások korösszetételét is. 1980 és 1996 között 12-ről 8,5 százalékra esett vissza a csak 30 évesnél fiatalabb személyekből álló háztartások aránya. A csak 60 évesekből vagy idősebbekből álló háztartások aránya viszont 18-ról 24 százalékra emelkedett. A háromgenerációs háztartások hányada 8-ról 5 százalékra esett vissza. Jellegzetes az egyszemélyes háztartások nem és kor szerinti összetétele: legalább 60 éves az egyszemélyes háztartások lakóinak 58 százaléka. Az egyszemélyes háztartások kétharmadában nő él egyedül, még hozzá az összes egyszemélyes háztartás 45 százalékában legalább 60 éves nő.

Az adatok szerint tehát a háztartások nagysága erősen összefügg nem- és korösszetételükkel. A háztartások azonos mintába kerülési esélyei mellett ez már önmagában előidézhethet reprezentativitási problémákat. Például vélhetően a nők a sokasághoz viszonyítva nagyobb számban kerülnek be a mintába, mivel nagy az egyszemélyes háztartások aránya, és ebben a háztartástípusban zömmel nők élnek. A fiatalok alulreprezentáltsága és az idősek nagyobb száma szintén háztartásszerkezeti jellemzőkre vezethető vissza, hiszen, mint láttuk, igen alacsony a csak fiatalokból álló háztartások aránya, míg minden negyedik háztartásban csak idősek élnek.

Fontos hangsúlyozni, hogy ez a hiba nem a Kish-kulcs sajátja, hasonló probléma jelentkezik minden más, háztartásmintán alapuló kétlépcsős mintavétel esetén. A probléma forrása ugyanis, mint láttuk, az egyenlőtlen kiválasztási valószínűségek megléte. Ez a hi-

ba jelentkeznek, ha azonos a háztartások bekerülési valószínűsége, és a háztartáson belül véletlen kiválasztást alkalmazunk. Ezért találtak egy 1990-es amerikai felmérésben azonos reprezentativitási problémákat mind a Kish-kulcs, mind a legutóbbi és a legközelebbi születésnapmódszer használata esetén (Kennedy; 1993). Ezért is szerepelhetnek a nők és főként az idős nők a tervezettnél nagyobb arányban a KSH legutóbbi születésnapmódszerrel készült mintájában (*Egészségi Állapot Felvétel...*; 1994).

Érdeemes rövid összehasonlítást végezni a mikrocenzus szerinti magyar és a Kish-kulcs létrehozásakor fennálló amerikai háztartásszerkezet között. Mint említettük, Kish a kiválasztási valószínűségek eltérését kiegyenlítő súlyozás elhagyhatósága mellett döntött (Kish; 1965). Állásfoglalását a súlyozott és a súlyozatlan becslések kis különbségére alapozta, a kis különbségek okaként pedig a korabeli háztartásszerkezet jellemzőit jelölte meg. Az általa hivatkozott háztartási adatok a következők.

5. tábla

Háztartások aránya felnőtt tagjaik szerint, 1957, Egyesült Államok

A felnőttek száma a háztartásban	1	2	3	4	5	6 vagy több
A háztartások százalékaránya	14,6	73,0	9,0	2,8	0,4	0,2

Kish értékelése szerint a kiválasztási valószínűségek szórása nagyon kicsi, mivel a háztartások zöme két felnőtt taggal rendelkezik, a többiek közül pedig gyakorlatilag mind 1, 3, vagy 4 felnőtt lakót számlál.

Ezzel szemben a hazai, 1996-os mikrocenzus szerint a magyar háztartások 26,1 százaléka egytagú. Ez csaknem kétszerese a Kish által tapasztalt aránynak.<sup>9</sup> Bár a nagy létszámú háztartások nálunk sem képviselik magukat jelentős számban, ez az eltérés önmagában is olyan mértékű, hogy óvatossá kell lenni a Kish-kulcs súlyozás nélküli honosításával.

### 3.2. A kulcs mint a reprezentativitási problémák forrása – a feltételezést alátámasztó példa

A fentiekben a reprezentativitási problémák egyik feltételezhető forrásaként jelöltük meg a kulcsot. Általában azonban nem választható el egymástól számszerűen a felmérésben alkalmazott két mintavételi lépcsőben, a háztartások, illetve a személyek kiválasztásakor fellépő hiba. Ezért szerencsés a következőkben leírt példa, ahol a rendelkezésre álló információk alapján jól elkülöníthető a két képcső hatása a minta létrejöttében.

Az ISSP Család II. felmérést a TÁRKI végezte 1994. május–júniusban. Ennek során a KSH 1990-es népszámlálásának lakásmintájából készült országos reprezentatív mintát használták, és a Kish-kulcsot alkalmazták a kérdezendő kiválasztására. A minta elem-

<sup>9</sup> Az Egyesült Államok háztartásszerkezete is jelentősen megváltozott az utóbbi évtizedekben. Növekvő háztartásszám és csökkenő átlagos háztartásnagyság volt megfigyelhető: 2000-ben az egyszemélyes háztartások aránya már 26 százalékos, ebből 9 százalékot tesznek ki a legalább 65 éves egyedül élők (www.census.gov). Az egyedül élők és az egyszemélyes háztartások arányának növekedése, illetve az együtt élő párok arányának csökkenése általános jelenség a mai modern társadalmakban (Bonvalet-Lelievre; 1997, de Jong; 1996).

száma 1500 fő. A háztartásokat véletlen (vagy legalábbis véletlent elérni kívánó) módon választották, majd minden háztartásból pontosan egy tagot jelöltek ki. Jól jellemezi a háztartásmintát használó kétlépcsős módszer kikerülhetetlen mintatorzító hatását már a minta háztartási struktúrája is. A mikrocenzus szerint a népesség 10,1 százaléka él egytagú háztartásban, ezzel szemben a Család II. mintájának 19,7 százaléka él egyedül.

Az alábbiakban a minta nem- és korösszetételét vizsgáljuk. A 6. tábla a minta korcsoport és nem szerinti százalékos megoszlását mutatja.

6. tábla

*A Család II. mintájának nem- és korösszetétele (n=1500)*

Korcsoport (éves)	Nem (százalék)		
	Férfi	Nő	Együtt
18–39	16,47	18,27	34,74
40–59	14,93	19,40	34,33
60 és idősebb	10,93	20,00	30,93
<i>Együtt</i>	<i>42,33</i>	<i>57,67</i>	<i>100,00</i>

A 7. táblában az 1994. év eleji sokasági arányokat láthatjuk.

7. tábla

*A magyarországi populáció 1994 elején (n=7 855 559)*

Korcsoport (éves)	Nem (százalék)		
	Férfi	Nő	Együtt
18–39	20,64	20,11	40,75
40–59	16,24	17,73	33,97
60 és idősebb	10,02	15,26	25,28
<i>Együtt</i>	<i>46,90</i>	<i>53,10</i>	<i>100,00</i>

*Forrás:* Magyar statisztikai évkönyv, 1993.

Összevetve az arányokat, szembeszökő a férfiak vártnál alacsonyabb számú előfordulása. A korcsoportokat nézve a középkorúak sokasági arányuknak megfelelően találhatók a mintában, míg a fiatalok alul-, az idősebbek felülreprezentáltak. A korcsoportok megoszlásának eltérése szignifikáns (khi-négyzet próbával vizsgálva, 0,000 szinten). A minta és a sokaság különbözősége a nemeket tekintve is szignifikáns (0,000 szinten). Ha nemcsak a marginálisok eltérését vizsgáljuk, hanem egyszerre a táblázat hat cellájának illeszkedését is,<sup>10</sup> megint csak szignifikáns eltérést kapunk (0,000 szinten).

A kulcs torzító hatását akkor lehetne érdemben elemezni, ha a kétlépcsős mintavétel lépcsőnként lenne jellemezhető. Ugyanis a fenti eltérések csak akkor tulajdoníthatók a végső lépcsőbeli kiválasztásnak, ha feltételezhetjük, hogy a háztartások egyforma valószínűséggel kerültek a mintába.

<sup>10</sup> A két nem és három korcsoport kategória kereszt-kombinációiból képzett hatértékű változó eloszlására végeztünk khi-négyzet tesztet.

Viszont megkísérelhető a kulcs okozta hatás vizsgálata, ha ismert a minta háztartásainak összes tagjából összeállított pszeudominta összetétele. Ekkor ugyanis ennek a pszeudomintának a nem- és korösszetételéhez viszonyíthatjuk a tényleges mintabeli megoszlásokat. Továbbá a pszeudominta és a sokasági adatok összevetésével a háztartásminta kiválasztási lépcsője is értékelhető.

A Család II. felvétel esetében a rendelkezésünkre állt nemcsak a kérdezettek, hanem a háztartások minden tagjának nemére és korára vonatkozó információ is. 47 esetben nem ismert valamelyik lakó neme vagy kora. Ezekkel a hiányokkal az eredeti 1500-as mintából 4397 használható személyi rekordhoz jutunk; a kérdezettek közül 3398-an töltötték be a 18-adik életévüket. A pszeudominta arányainak százalékos megoszlását a 8. tábla mutatja.

8. tábla

*A Család II pszeudomintájának nem- és korösszetétele (n=3389)*

Korcsoport (éves)	Nem (százalék)		
	Férfi	Nő	Együtt
18–39	19,50	19,92	39,42
40–59	16,50	19,21	35,71
60 és idősebb	10,00	14,87	24,87
<i>Együtt</i>	<i>46,00</i>	<i>54,00</i>	<i>100,00</i>

A pszeudominta és az eredeti személyi minta összehasonlításában a tendenciák egyeznek az előbbi összehasonlításkor kapottal: a személyi mintában felülreprezentáltak az idősek, míg a fiatal korosztály alulreprezentált. Az eltérés szignifikáns (0,000 szinten). Az eltérés a nemeket tekintve is szignifikáns (0,004 szinten). A marginálisokon kívül a hat kategória illeszkedése is szignifikáns eltérést mutat (0,000 szinten).

A két összehasonlítás tapasztalatai egyeznek, mert a pszeudominta igen jól illeszkedik a sokasághoz. A pszeudominta és a sokaság összehasonlításában most is alulreprezentáltak a férfiak, de az eltérés nem szignifikáns ( $p=0,31$ ). A korcsoportok eltérése sem szignifikáns ( $p=0,09$ ), és a cellákat vizsgálva a hat kategória illeszkedése is elfogadható ( $p=0,222$ ).

Eredményeink lehetővé teszik a mintaválasztás egymástól szétválasztott lépcsőfokainak értékelését. A pszeudominta jól illeszkedik a várható arányokhoz. Eszerint a háztartásválasztás véletlen volta nem cáfolható, a háztartásminta még reprezentatívnak minősíthető. A személyválasztási eljárást, azaz a kulcsot a hiba okai között önálló tényezőként különítettük el.

### 3.3. A kulcs mint a reprezentativitási problémák forrása – a feltételezés bizonyítása

A fenti eredményt megkérdőjelezheti az, hogy csak egy mintát vizsgáltunk, és annak összetétele akár a véletlen egyszeri játékkal is magyarázható. Tegyük most ezt az értékelést jobban általánosíthatóvá.

Elemzésünkben a TÁRKI adatbázisát, a Magyar Háztartási Panel IV. hullámfelmérést használtuk. A kutatást vezető intézmények a BKE Szociológiai Tanszék és a TÁRKI. Az

adatfelvétel időpontja 1995. Módszere: kérdőíves adatfelvétel. Esetszám: 8043. Ez országos háztartási minta, 6306 személy adatait tartalmazza a bázis. Kizártuk az elemzésből azokat a személyeket, akiknek hiányzott az információ a koráról vagy a neméről. A többiek közül 4188-an töltötték be 1994-re a 18-adik életévüket. Így 4188 személy – 1986 háztartás – adatait használtuk fel. 813 esetben hiányzott az információ valamely háztartás tag neméről vagy koráról, de elsősorban nem válaszmegtagadás miatt, nem vettek fel adatokat a 16 éven aluliakról, a háztartásból kivált, de még nyilvántartott személyekről stb.

Az elemzés során mintavételi szimulációt végzünk, azaz úgy kezeljük a mintát, mintha egy virtuális sokaság lenne. A továbbiakban a minta megnevezést is pszeudosokaságra váltjuk. Ezen a pszeudosokaságon teszteljük a Kish-kulcs használatának hatását a minták nem- és korösszetételére.

9. tábla

*A Háztartás Panelből kapott sokaság összetétele (n=4188)*

Korcsoport (éves)	Nem (százalék)		
	Férfi	Nő	Együtt
18–39	19,79	20,51	40,30
40–59	14,66	17,60	32,26
60 és idősebb	10,94	16,50	27,44
<i>Együtt</i>	<i>45,39</i>	<i>54,61</i>	<i>100,00</i>

Kiszűrendő a mintaösszetételnek a háztartásválasztás által okozott ingadozását, a kulcs teljesítményének értékelését nem egyetlen minta kiválasztásával értékeltük. Ehelyett a várható minta összetételét számítottuk ki. A várható minta összetételét a következőképpen kaphatjuk: egy adott nem/korcsoport kategória mintabeli arányának mint valószínűségi változónak a várható értéke adja a kategória várható mintabeli arányát (részletesebben lásd alább az /1/ egyenletet).

10. tábla

*Várható mintaösszetétel Kish-kulcs alkalmazásakor*

Korcsoport (éves)	Nem (százalék)		
	Férfi	Nő	Együtt
18–39	17,27	19,47	36,74
40–59	12,92	17,04	29,96
60 és idősebb	11,87	21,43	33,30
<i>Együtt</i>	<i>42,06</i>	<i>57,94</i>	<i>100,00</i>

Megfigyelhetjük az idős korosztály felülreprezentáltságát, leginkább az idős nőket. A másik két korosztály a kellőnél kisebb számban szerepel, főként a fiatal és a középkorú férfiak. A férfiak a vártól kisebb számarányban képviseltetik magukat. Az eltérések megegyeznek a valós felmérések tapasztalataival.

Eredményeink szerint tehát szisztematikus torzítás lép fel a Kish-kulcsot alkalmazó mintavételben. Ezt a torzítást maga a mintavétel módja okozza: az, hogy minden háztartásból a véletlent imitáló módon, azonos valószínűségekkel választunk. Nem a kulcs okozza a hibát, ugyanez a probléma lép fel például a legutóbbi születésnap módszerének alkalmazásánál is. Ugyanakkor a kulcs esete azért szerencsés, mert használatakor ellenőrizhető a minta nem és kor szerinti összetétele, azaz esetleg javítható is.

#### 4. A KISH-KULCS MÓDOSÍTÁSA

Az alábbiakban a kulcs módosítására teszünk kísérletet, olyan céllal, hogy a minta a nem/korcsoport jellemzők szerint minél közelebb legyen a sokasági paraméterekhez.

A Kish-kulcs módosításakor a mintavételi eljárás alapvető vonásait állandónak tételjük fel. Így nem változtatunk azon, hogy a háztartások azonos valószínűséggel kerülnek a mintába, s hogy minden háztartásból pontosan egy személyt választunk ki. A kulcsból azt a jellemzőt is megtartjuk, hogy a háztartás tagjait nemük és koruk szerint sorba tesszük, és a kérdőívhez csatolt táblából keresheti ki a kérdező, hogy a sor hányadik tagját kérdezze. A módosítás csak ezekre a kiválasztási táblákra terjed ki.

Felhasználva a sokaság háztartásszerkezetének ismeretét, a kiválasztási táblákból egyértelműen meghatározhatjuk a minta várható nem- és korcsoportszerkezetét. A feladat tehát olyan táblák elkészítése, amelyek használata esetén a minta nem és korcsoport szerinti összetétele megegyezik, vagy a „lehető legközelebb van” a sokasági megoszláshoz. A feladat még egy paramétert tartalmaz, ez a kérdőívekhez csatolt táblák típusainak száma. Minél több táblafajtát használhatunk, annál finomabb valószínűség-eloszlás jön létre, s vélhetően annál közelebb hozható a minta a sokasághoz. A Kish-kulcs eredetije 8 táblát használ fel, és ezeket különböző arányban osztja ki a kérdőívek között, 1/12-ekben adva meg a kiosztási arányt. Az alábbi módosításban 12-ben rögzítjük a táblatípusok számát, míg az egyes táblatípusok kiosztási arányát továbbra is 1/12-ekben adjuk meg. A meghatározásra váró ismeretlenek tehát a 3. táblához hasonló tábla elemei,  $12 \times 6$  érték.

Az ismeretlenek segítségével egyértelműen meghatározható, hogy egy  $i$  létszámú háztartásnak a nem/kor szerinti sorban  $j$ -edik tagja mekkora valószínűséggel kerül kiválasztásra, feltéve, hogy háztartása már bekerült a mintába. Ha  $j > 6$ , ez a valószínűség 0. Más esetben ez a valószínűség 12-ed része annak, ahányszor a keresett táblázat  $i$ -edik oszlopában a  $j$  szerepel. Tehát a kulcs módosításának feladata egyszerűsíthető: nem a táblákat kell meghatározni, csak azokat a valószínűségeket (12-edekben), amekkora eséllyel a táblák használatával az  $i$  tagú háztartások  $j$ -edik tagja kiválasztásra kerül. Jelöljük a meghatározandó feltételes valószínűségeket  $p_{ij}$ -vel ( $i$  tagú háztartás sorban  $j$ -edik tagjához tartozó kiválasztási valószínűség, miután a háztartás már bekerült a mintába). A 12-edekben adott  $p_{ij}$  valószínűségek segítségével a 12 tábla már egyértelműen meghatározható; például  $p_{ij}=5/12$  esetén a keresett táblázat  $i$ -edik oszlopában a  $j$ -t 5-ször kell szerepeltetni (az oszlop tetszőleges celláiban). Az ismeretlenek tehát:  $p_{21}, p_{22}, p_{31}, p_{32}, p_{33}, \dots, p_{61}, p_{62}, p_{63}, p_{64}, p_{65}, p_{66}$ . (Értelmezhető  $p_{11}$  is, ennek értéke nyilván 1.)

Az  $i$  tagú háztartásban lakó  $j$ -edik személy kiválasztási valószínűsége a háztartás kiválasztási valószínűségének és az egyén háztartáson belüli feltételes kiválasztási valószínűségének szorzata. Az utóbbi érték a  $p_{ij}$ -vel jelölt ismeretlen paraméter, az előbbi érték viszont a sokaság háztartásszerkezetének függvénye. Mivel a háztartások azonos valószínűségi



nűséggel kerülnek a mintába, az  $i$  létszámú háztartás kiválasztási valószínűsége az  $i$  tagú háztartások sokasági arányával egyezik meg. Jelölje ezt az értéket  $H_i$ , a feladat input paramétere tehát:  $H_1, \dots, H_6$ .

A feladat megoldásához szükség van annak valószínűségére, hogy egy  $i$  tagú háztartásban  $j$ . tagja fiatal férfi, fiatal nő, középkorú férfi, középkorú nő, idős férfi, illetve idős nő. Ez  $3 \times 2$  érték, amely egy  $3 \times 2$  dimenziós mátrixszal adható meg. A fentiekén kívül a feladat input paramétere tehát összesen még 21 darab mátrix, jelölve:  $\mathbf{a}_{11}, \mathbf{a}_{21}, \dots, \mathbf{a}_{66}$ . Például  $a_{21}[11]$  jelöli annak a valószínűségét, hogy egy 2 tagú háztartás sorban első tagja fiatal férfi,  $a_{21}[21]$  annak a valószínűségét, hogy ugyanő középkorú férfi,  $a_{21}[12]$  pedig annak valószínűségét, hogy fiatal nő.

Most már megadható a mintában várható adott nemű, adott korcsoportoz tartozó személyek aránya, és ez a hat érték az előbbi módon ismét egy  $3 \times 2$  dimenziós mátrixba rendezhető. Jelölje a mátrixot  $\mathbf{a}$ .  $a[11]$  a fiatal férfiak arányát jelöli,  $a[21]$  a középkorú férfiak arányát stb. Ezek az értékek kifejezhetők a fenti paraméterek függvényében:

$$a[ij] = \sum_{k=1..6} H_k \left( \sum_{l=1..k} p_{kl} a_{kl}[ij] \right) \quad i = 1, 2, 3 \quad j = 1, 2. \quad /1/$$

A célként jelölt sokasági nem/korcsoport megoszlás szintén egy  $3 \times 2$  dimenziós mátrixszal adható meg. A mátrixot jelölje  $\mathbf{A}$ , cellái a mintabeli megoszlást jelölő  $a[ij]$ -khez hasonlóan értelmezhetők. Így például  $A[11]$  a fiatal férfiak sokasági arányával egyezik meg. A feladat szerint keressük  $p_{ij}$ -k függvényében azon  $a[ij]$  értékeket, amelyeknek értéke a megfelelő  $A[ij]$  értékkel egyezik meg, ekkor a minta nem/korcsoport megoszlása pontosan reprodukálja a sokaságét.

A feladat megfogalmazása tehát: adott az  $\mathbf{A}$  mátrix, adottak az  $\mathbf{a}_{21}, \dots, \mathbf{a}_{66}$  mátrixok. Keressük azon  $p_{21}, \dots, p_{66}$  skalárokat, amelyek kielégítik az alábbi egyenletet:

$$\sum_{i=1,2,3} \sum_{j=1,2} a[ij] - A[ij] = 0, \quad /2/$$

az alábbi feltételekkel:

$$\begin{aligned} \sum_{j=1..i} p_{ij} &= 1 \quad \forall i\text{-re}, \\ p_{ij} &> 0 \quad \forall i, j\text{-re}, \\ p_{ij} &= k_{ij} / 12 \quad \forall i, j\text{-re, ahol } k_{ij} \text{ egész.} \end{aligned}$$

0-nál nagyobb kiválasztási valószínűségeket kívánunk meg, ugyanis a mintaválasztási eljárásokkal szembeni alapvető követelmények közé tartozik a pozitív kiválasztási valószínűségek megléte. A sokaság megváltoztatását jelentené a 0 értékű kiválasztási esély megengedése, hiszen ezzel szűkítenénk a kiválaszthatók körét.

A Microsoft Excel Solver bővítményprogramját használjuk a feladat megoldásának keresésére.<sup>11</sup> A feladat egy többváltozós nemlineáris egyenlet egészértékű megoldását keresi, korlátozó feltételek mellett.

<sup>11</sup> A Microsoft Excel Solver az általános redukált gradiens (GRG2) nemlineáris optimalizálási eljárást használja. A lineáris és az egész értékű problémákra a változókat korlátozó szimplex módszert, valamint az elágazás és korlátozás (branch-and-bound) módszert alkalmazza.

Az input paraméterek megadásakor a korábban is használt Magyar Háztartás Panel mintáját szerepeltetjük sokaságként. A sokasági nem/korcsoport arányok, az  $A[ij]$ -k értéke a 9. táblában szerepel.

Az egyenletnek nincs olyan megoldása, amely a feltételeket kielégítené.

Felvethető az a kérdés, hogy megoldható-e a probléma, ha nem köt minket a táblatípusok számának korlátja. A táblák számának növelése nyilván költségnövekedést eredményez, nagyobb adminisztrációt, nagyobb nyomdaköltséget igényelhet. Azonkívül a minta nagysága maga is felső korlátja a kiosztható táblák számának. Az elvi probléma azonban végiggondolásra érdemes. A probléma megfogalmazásakor ekkor a kiválasztási valószínűségekre vonatkozó feltételek közül elhagytuk az egészértékűségi feltételt, és a pozitivitás feltétel helyett  $1/100$ -os alsó határt adtunk meg.

A feladatnak azonban ilyenkor sincs megoldása.

Tehát nem reprodukálható pontosan a sokaság nem/korcsoport megoszlása. Ehelyett kereshetjük azokat a táblákat, amelyek használata esetén a sokasághoz „legközelebb levő” mintát kapunk. A legközelebb levő fogalmának operacionalizálásához a távolság fogalmát is meg kell határozni. A távolságot kétféleképpen határozhatjuk meg. Elsőként a sokaság és a minta távolságát a khi-négyzet statisztikához hasonlóan határozzuk meg: a mintából várt és a sokasági arányok négyzetes eltéréseinek és a várt arányok a hányadosát összegeztük a hat cellára:

$$f(a) := \sum_{i=1,2,3} \sum_{j=1,2} (a[ij] - A[ij])^2 / A[ij]. \quad /3/$$

Az optimális táblák esetén ez a távolság minimális, azaz most a feladat megfogalmazása:

$$f(a) \rightarrow \min,$$

a korábbi feltételekkel:

$$\begin{aligned} \sum_{j=1\dots i} p_{ij} &= 1 \quad \forall i\text{-re}, \\ p_{ij} &> 0 \quad \forall i, j\text{-re}, \\ p_{ij} &= k_{ij} / 12 \quad \forall i, j\text{-re, ahol } k_{ij} \text{ egész.} \end{aligned}$$

Ez egy nemlineáris egészértékű optimalizálási probléma, megszorító feltételekkel.

Ugyanakkor más megközelítésben is definiálhatjuk a távolság fogalmát. A másik távolságfüggvény alkalmazása a felmérések elemzésekor gyakran alkalmazott súlyozási eljárással kapcsolatban merül fel. A súlyozás célja a torzítások csökkentése. A súlyok alkalmazásakor viszont gyakran megnő a becslések varianciája. (Lásd a 6. lábjegyzetet.) A Kish-kulcs módosításakor a nem és kor szerinti reprezentativitást tartottuk szem előtt. A minta megoszlásának eltérése a kívánatostól súlyozással korrigálható, ez az utólagos rétegzés a minta nem és korcsoport arányait illeszti a sokaságéhoz. A variancia növekedése ebben az esetben a súlyok változékonyságával hozható összefüggésbe. Pontosabban, a variancianövekedés a súlyok négyzetösszegének monoton függvénye. Így adódik a kulcs újabb szempont szerinti módosítása: keressük azt az eljárást, amely a minimális

négyzetösszegű súlyokat eredményezi (a súlyok összege itt állandó: 1). A korábbi jelölésekkel a súlyok négyzetösszege:

$$\begin{aligned}\sum_{k=1\dots n} W_k^2 &= \sum_{i=1,2,3} \sum_{j=1,2} \sum_{k:\text{korcsoport}=i \text{ nem}=j} W_k^2 \\ &= \sum_{i=1,2,3} \sum_{j=1,2} \sum_{k:\text{korcsoport}=i \text{ nem}=j} A[ij]^2 / a[ij]^2 \\ &= \sum_{i=1,2,3} \sum_{j=1,2} n_{ij} A[ij]^2 / a[ij]^2 = \sum_{i=1,2,3} \sum_{j=1,2} n_{ij} A[ij]^2 / (n_{ij} / n)^2 \\ &= n \sum_{i=1,2,3} \sum_{j=1,2} A[ij]^2 / (n_{ij} / n) = n \sum_{i=1,2,3} \sum_{j=1,2} A[ij]^2 / a[ij],\end{aligned}$$

ahol  $n$  a mintanagyság,  $n_{ij}$  az  $i$ -edik korcsoporthoz és a  $j$ -edik nemhez tartozók száma a mintában.

Mint látható, a súlyok négyzetösszege függ a mintanagyságtól. Ez a függés kiküszöbölhető, ha a mintanagyságot rögzítettnek tekintjük. Ekkor az optimumkeresés során a négyzetösszeg helyett a

$$g(a) := \sum_{i=1,2,3} \sum_{j=1,2} A[ij]^2 / a[ij] = (1/n) \sum_{k=1\dots n} W_k^2 \quad /4/$$

függvényt optimalizáljuk. A korlátozó feltételek ugyanazok, mint fent, az  $f$  függvény használata esetén.

A 10. táblában látható, hogy a Kish-kulcs alkalmazása esetén milyen  $a[ij]$ -ket kapunk, azaz milyen a mintabeli várható aránya az adott nemű, adott korcsoporthoz tartozó személyeknek. Ezeket az értékeket  $f$ -be helyettesítve a függvény értéke 0,021573301. Ha találunk olyan paraméterértékeket, ahol az  $f$  ennél kisebb értéket vesz fel, akkor a Kish-kulcsnál jobb nem/korcsoport illeszkedést produkáló kiválasztási eljárást határozhatunk meg.

Hasonlóan, a  $g$ -be helyettesítve a Kish-kulcs alkalmazásakor kapott  $a[ij]$ -ket 1,018901199-et kapunk. Ennél kisebb értéket adó paraméterértékek megtalálása esetén olyan kiválasztási eljárást határozhatunk meg, amely az utólagos súlyozás szempontjából kedvezőbb mintát produkál.

A függvényeknek létezik elvi optimuma a megadott tartományon.  $f$  és  $g$  ugyanis folytonos függvény, a megadott tartomány zárt és korlátos. *Weierstrass* tétele szerint a folytonos függvénynek van legkisebb és legnagyobb értéke korlátos és zárt tartományon. Az elvi optimum megtalálása azonban nem egyszerű matematikai probléma. Speciális esetektől eltekintve a nemlineáris optimalizálási feladatok olyan algoritmussal oldhatók meg, mellyel általában nem garantálható, hogy a talált optimum globális és nem csupán lokális szélsőérték. A megoldás keresésére ismét az Excel Solver programját használjuk.

Mindkét feladatot megoldjuk az egészértékűségi feltétel mellett (tehát 12 tábla használatával) és annak elhagyásával is. A 11. tábla foglalja össze az eredményeket.

Az a mátrixokat összehasonlítva megállapíthatjuk, hogy a négy variáns gyakorlatilag egyező eredményt ad: az értékek első két tizedes jegyre kerekítve megegyeznek a négy esetben. Eszerint nem ad jobb eredményt az egészértékűségi feltétel elhagyása, tehát nem érdemes 12-nél több táblával dolgozni.

Összevetve a 9. táblabeli sokasági nem/korcsoport megoszlást az eredményül kapott  $a[ij]$ -kel, láthatóan javul a Kish-kulcshoz képest a fiatal korosztályok illeszkedése,

ugyanígy javul az idős nők aránya. Viszont a középkorú nők és az idős férfiak részaránya kissé távolabb kerül a sokaságtól, mint a Kish-kulcs esetén.

Az eredményül kapott  $p_{ij}$  értékeket értelmezve elmondható, hogy a várt tendenciák tapasztalhatók. A kétszemélyes háztartásokból a Kish-kulcsban adottnál nagyobb valószínűséggel választjuk az első, várhatóan inkább férfi tagot, a nők Kish-kulcsos felülreprezentáltságát korrigálva. Ugyanígy a három- és négytagú háztartásokból sorban utolsó, várhatóan fiatal lány tagot választjuk nagyobb valószínűséggel, korábbi alulreprezentáltságukat javítva.

11. tábla

## Optimalizálási eredmények

Eredeti Kish-kulcs									
Függvényértékek a Kish-kulcs behelyettesítésekor az $f$ értéke: 0,021573301 a $g$ értéke: 1,018901199	Várható nem/kor eloszlás (a mátrix)		$p_{ij}$ -k						
	0,1727	0,1947	$p_{21}$	1/2	$p_{31}$	1/3	$p_{41}$	1/4	
	0,1292	0,1704	$p_{22}$	1/2	$p_{32}$	1/3	$p_{42}$	1/4	
	0,1187	0,2143			$p_{33}$	1/3	$p_{43}$	1/4	
						$p_{44}$	1/4		
			$p_{51}$	1/6	$p_{61}$	1/6			
			$p_{52}$	1/6	$p_{62}$	1/6			
			$p_{53}$	1/4	$p_{63}$	1/6			
			$p_{54}$	1/6	$p_{64}$	1/6			
			$p_{55}$	1/4	$p_{65}$	1/6			
					$p_{66}$	1/6			
$f$ optimalizálása									
Korlátozó feltételek:	optimum	Az optimummal kapott várható nem/kor eloszlás (a mátrix)		Optimumhely ( $p_{ij}$ -k)					
$\sum_{j=1}^6 p_{ij} = 1 \quad \forall i$ -re $p_{ij} > 0 \quad \forall i, j$ -re $p_{ij} = k_{ij}/12 \quad \forall i, j$ -re, ahol $k_{ij}$ egész szám.	0,013938589	0,1856	0,1994	$p_{21}$	2/3	$p_{31}$	2/12	$p_{41}$	1/12
		0,1281	0,1616	$p_{22}$	1/3	$p_{32}$	1/12	$p_{42}$	3/12
		0,1311	0,1942			$p_{33}$	9/12	$p_{43}$	1/12
								$p_{44}$	7/12
				$p_{51}$	1/12	$p_{61}$	1/12		
				$p_{52}$	8/12	$p_{62}$	7/12		
				$p_{53}$	1/12	$p_{63}$	1/12		
				$p_{54}$	1/12	$p_{64}$	1/12		
		$p_{55}$	1/12	$p_{65}$	1/12				
				$p_{66}$	1/12				
$\sum_{j=1}^6 p_{ij} = 1 \quad \forall i$ -re $p_{ij} > 0 \quad \forall i, j$ -re	0,013216638	0,1875	0,2015	$p_{21}$	0,6539	$p_{31}$	0,2351	$p_{41}$	0,0100
		0,1293	0,1584	$p_{22}$	0,3461	$p_{32}$	0,0100	$p_{42}$	0,3728
		0,1307	0,1926			$p_{33}$	0,7549	$p_{43}$	0,0100
								$p_{44}$	0,6072
				$p_{51}$	0,0100	$p_{61}$	0,0100		
				$p_{52}$	0,8877	$p_{62}$	0,9500		
				$p_{53}$	0,0100	$p_{63}$	0,0100		
				$p_{54}$	0,0823	$p_{64}$	0,0100		
		$p_{55}$	0,0100	$p_{65}$	0,0100				
				$p_{66}$	0,0100				

(A tábla folytatása a következő oldalon.)

(Folytatás.)

g optimalizálása									
Korlátozó feltételek:	Optimum	Az optimummal kapott várható nem/kor-eloszlás (a mátrix)		Optimumhely ( $p_{ij}$ -k)					
$\sum_{j=1}^n p_{ij} = 1 \quad \forall i$ -re $p_{ij} > 0, \quad \forall i, j$ -re, $p_{ij} = k_{ij}/12 \quad \forall i, j$ -re, ahol $k_{ij}$ egész szám.	1,012869186	0,1854	0,1960	$p_{21}$	2/3	$p_{31}$	2/12	$p_{41}$	2/12
		0,1314	0,1607	$p_{22}$	1/3	$p_{32}$	1/12	$p_{42}$	3/12
		0,1323	0,1942			$p_{33}$	9/12	$p_{43}$	1/12
						$p_{44}$	6/12		
				$p_{51}$	1/12	$p_{61}$	1/12		
				$p_{52}$	7/12	$p_{62}$	7/12		
				$p_{53}$	1/12	$p_{63}$	1/12		
				$p_{54}$	2/12	$p_{64}$	1/12		
				$p_{55}$	1/12	$p_{65}$	1/12		
						$p_{66}$	1/12		
$\sum_{j=1}^n p_{ij} = 1 \quad \forall i$ -re $p_{ij} > 0, \quad \forall i, j$ -re	1,012269154	0,1863	0,1993	$p_{21}$	0,6574	$p_{31}$	0,2543	$p_{41}$	0,0270
		0,1311	0,1591	$p_{22}$	0,3426	$p_{32}$	0,0100	$p_{42}$	0,3886
		0,1323	0,1919			$p_{33}$	0,7357	$p_{43}$	0,0100
								$p_{44}$	0,5744
				$p_{51}$	0,0100	$p_{61}$	0,0100		
				$p_{52}$	0,5594	$p_{62}$	0,7638		
				$p_{53}$	0,0100	$p_{63}$	0,1962		
				$p_{54}$	0,4106	$p_{64}$	0,0100		
				$p_{55}$	0,0100	$p_{65}$	0,0100		
						$p_{66}$	0,0100		

A két függvény által az egészértékűségi feltétel mellett, illetve annak elhagyásával kapott eljárások nem különböznek lényegesen a  $p_{ij}$ -ket tekintve, és gyakorlatilag azonos eredményt adnak a várható mintaösszetétel szerint is. Továbbá, ha  $f$  optimumát  $g$ -be helyettesítjük, akkor  $g$  optimumához igen közeli értéket kapunk, és fordítva. Ez azt jelenti, hogy az optimális megoldások mindkét szempont szerint megfelelők. Mivel nem különböznek lényegesen a megoldások, válasszuk az  $f$  által az egészértékűségi feltétel mellett adott optimumot. A kapott  $p_{ij}$  értékhez a következő Kish-tábla illeszthető.

12. tábla

Módosított Kish-tábla

A tábla aránya	A tábla jelzete	A kiválasztott sorszáma, ha a felnőttek száma					
		1	2	3	4	5	6 vagy több
1/12	1.	1	1	1	1	1	1
1/12	2.	1	1	1	2	2	2
1/12	3.	1	1	2	2	2	2
1/12	4.	1	1	3	2	2	2
1/12	5.	1	1	3	3	2	2
1/12	6.	1	1	3	4	2	2
1/12	7.	1	1	3	4	2	2
1/12	8.	1	1	3	4	2	2
1/12	9.	1	2	3	4	2	3
1/12	10.	1	2	3	4	3	4
1/12	11.	1	2	3	4	4	5
1/12	12.	1	2	3	4	5	6

A 12. tábla közvetlenül használható a felmérésekben.

\*

Legfontosabb megállapításaink a következők.

- A kulccsal készített magyarországi minták nem és kor szerinti jellemzők alapján lényegesen eltérnek a sokaságtól. Ennek elvi oka az, hogy kétlépcsős eljárásban, háztartási mintát alkalmazva a végső, személyi minta az aktuális háztartásszerkezet függvénye.
- A kulcs eredményesen módosítható, ha célunk a minta adott jellemzők szerinti reprezentativitásának javítása.

Sikerült bizonyítanunk a kulccsal kapott minta és az aktuális sokasági háztartásszerkezet kapcsolatát. Az általunk felkutatott források nem bizonyították ezt a kapcsolatot, sőt, általában nem is említették ezt a tényezőt. Másik eredményünk, a kulcs módosítása önmagában is értékelhető. De fontos szempont, hogy a megoldásra alkalmazott algoritmus sokkal szélesebb körben is alkalmazható. Használatával meghatározhatunk más, nem a kor és nem jellemzők szerinti reprezentativitást javító optimális kulcsvariánsokat is.

Természetesen felvetődnek további, még nyitva álló kérdések. Elemzésünket a nem és kor szerinti reprezentativitásra korlátoztuk. Vizsgálandó lenne a Kish-kulccsal kapott minták más, lényeges jellemzők szerinti illeszkedése. A módosítás eredményeként kapott eljárás tesztelése is szükséges lenne, ugyanis a nem és kor szerinti illeszkedés javulása nem feltétlenül jelenti a minta más jellemzők szerinti javulását. A módosított kulcs használatával megváltozott mintavételi valószínűségek is további vizsgálatokra szorulnak. Továbbá a kulcs módosítását csupán a Magyar Háztartás Panelből vezettük le, a megnyugtató érvényességhez szükséges lenne ugyanezt az optimalizáló feladatot a teljes magyar háztartásstruktúra jellemzőiből kiindulva, népszámlálási adatokon végrehajtani.

#### FORRÁS- ÉS IRODALOMJEGYZÉK

- BINSON, D. – CANCHOLA, J. A. – CATANIA, J. A. (2000): Random selection in a telephone survey: A Comparison of the Kish, Next-Birthday, and Last-Birthday Methods. *Journal of Official Statistics*, 16. évf. 1. sz. 53–59. old.
- BONVALET, C. – LELIEVRE, E. (1997): The transformation of housing and household structures in France and Great Britain. *International Journal of Population Geography*, 3. évf. 3. sz. 183–201.
- Egészségi Állapot Felvétel (1994) – Életmód, kockázati tényezők (1996)*. Központi Statisztikai Hivatal, Budapest.
- GROVES, R. M. (1989): *Survey errors and survey costs*. John Wiley and Sons, Inc., New York.
- GROVES, R.M. – BIEMER, P. P. – LYBERG, L. E. – MASSEY, J. T. – NICHOLLS, W. L. – WAKSBERG, J. (szerk.) (1988): *Telephone survey methodology*. John Wiley and Sons, Inc., New York.
- DE JONG, A. H. (1996): National household forecasts 1996: fewer and fewer couples are married. *Maandstatistiek van de Bevolking*, 45. évf. 5. sz. 18–27. old.
- KENNEDY, J. M. (1993): *A comparison of telephone survey respondent selection procedures*. <http://www.indiana.edu/~csr/aapor93.html>
- KISH, L. (1949): A procedure for objective respondent selection within the household. *Journal of the American Statistical Association*, 44. évf. 380–387. old.
- KISH, L. (1965): *Survey sampling*. John Wiley and Sons, Inc., New York.
- LAVRAKAS, P. J. (1993): *Telephone survey methods: Sampling, selection and supervision*. Applied Social Research Methods. 7. köt.
- LEVY, P. S. – LEMESHOW, S. (1999): *Sampling of populations*. John Wiley and Sons, Inc. New York.
- Magyar Statisztikai Évkönyv 1993 (1994)*. Központi Statisztikai Hivatal, Budapest.
- Magyar Statisztikai Évkönyv 1999 (2000)*. Központi Statisztikai Hivatal, Budapest.
- Market research methods*. University of Southampton. <http://ispstat.alcd.soton.ac.uk/am306/quant5.txt>
- Mikrocenzus, 1996. Főbb eredmények (1996)*. Központi Statisztikai Hivatal, Budapest.
- Mikrocenzus, 1996. A népesség és a lakások jellemzői (1996)*. Központi Statisztikai Hivatal, Budapest.
- OLDENDICK, R. W. – BISHOP, G. G. – SORENSON, S. B. – TUCHFARBER, A. J. (1988): A comparison of the Kish and Last Birthday Methods of respondent selection in telephone surveys. *Journal of Official Statistics*, 4. évf. 4. sz. 307–318. old.

- PIAZZA, T.: *Respondent selection for CATI/CAPI (Equivalent to the Kish method)*. <http://srcweb.berkeley.edu:4229/res/rsel.html>  
*SCSU Survey, 1997. Minnesota State Lottery*. <http://www.lottery.state.mn.us/gambling/stcloud.html>  
*Táblaképek az egészségről – A veresegyházi példa (2001)*. MTA Szociológiai Kutatóintézet – Fekete Sas Kiadó, Budapest.
- VOSS, D. S. – GELMAN, A. – KING, G.: *The Polls – A review*. *Preelection survey methodology: Details from eight polling organizations, 1998 and 1992. Public Opinion Quarterly*, 59. évf. 98–132. old.
- TARJÁNYI J.: *Módszertani problémák a telefonos közvélemény-kutatásokban*. <http://www.c3.hu/scripta/scripta0/replika/1920/14vita.htm>

## SUMMARY

The problem of drawing a person from a household often occurs at the final stage of a survey design e. g. in telephone surveys, after contacting the households using random digit dialing. The Kish grid offers an algorithm for this random selection. The authors found that in contrast with the widely held opinion, the grid is not capable of presenting the population by gender and age. This misconception stems from the fact that when the Kish grid was developed in the 1950's, both randomness and representativity could be achieved through the method due to the household structure in the USA. On the basis of their calculation, this does not hold for the recent Hungary. Finally, the authors suggest a modification to the Kish grid that is more appropriate for selecting a representative sample.

# FÜGGETLEN MINTÁK ÖSSZEHASONLÍTÁSA ÚJ RANGSOROLÁSOS ELJÁRÁSOKKAL\*

VARGHA ANDRÁS

A társadalomtudományokban gyakran találkozunk olyan nem kvantitatív változókkal, amelyek értékskálája csupán az ordinalitás kritériumának tesz eleget. Ilyen változók esetén a nagyság szintek átlag segítségével történő összehasonlítása vitatható.

A jelen tanulmány alternatív megoldásként hívja fel a figyelmet a sztochasztikus összehasonlításra, amelynek kulcsfogalmai a sztochasztikus egyenlőség (SZTE) és a sztochasztikus homogenitás (SZTH). Két sokaságot akkor mondunk sztochasztikusan egyenlőnek valamely  $X$  változó tekintetében, ha  $P(X_1 > X_2) = P(X_1 < X_2)$  teljesül, ahol  $X_1$  és  $X_2$  a két sokaságból egymástól függetlenül kiválasztott véletlen megfigyelés. Az SZTH az SZTE általánosítása, segítségével kettőnél több sokaság sztochasztikus ugyanakkoraságát definiáljuk.

Tanulmányunkban az SZTH tesztelésével kapcsolatban 15 statisztikai próba előnyeit és hátrányait vetjük össze számítógépes szimulációs elemzés segítségével.

TÁRGYSZÓ: Sztochasztikus homogenitás. Rangsorolós csoport-összehasonlítás.

A társadalomtudományi kutatások egyik leggyakoribb kérdése, hogy valamely  $X$  változó értékei általában ugyanakkorák-e bizonyos csoportokban. Illusztrációul felsorolunk néhány ilyen jellegű szakmai kérdést.

- A magyarországi felnőtt lakosságon belül az olvasásra szánt idő (mondjuk 1 hétre vonatkoztatva) ugyanakkora-e különböző iskolázottsági szinteken?
- Független-e az ARIEL mosópor ötfokú szimpátiaskálán mért kedveltségi szintje attól, hogy a megkérdezett személy a fővárosban, vidéki városban vagy községben lakik?
- Van-e különbség a 19 magyar megye között a tekintetben, hogy lakosai mennyire vannak megelégedve a kormánykoalícióval, ha a megelégedettséget ismét ötfokú skálán mérjük?
- Ugyanolyan hatékony-e valamely A, B és C terápiás eljárás a dohányzás tekintetében, ha a hatékonyságot azzal mérjük, hogy a terápiában részt vett személy átlagosan hány cigarettát szív el naponta a terápia befejezése után 3 hónappal?

Ha  $X$  kvantitatív változó, akkor az ilyen típusú kérdésekre megfelelő statisztikai válasz a

$$H_0: \mu_1 = \mu_2 = \dots = \mu_h$$

nullhipotézis vizsgálata, ahol  $\mu_1, \mu_2, \dots, \mu_h$  az  $X$  változó elméleti átlaga (várható értéke)

\* A tanulmány megírásához segítséget nyújtott a T032157 sz. OTKA-pályázat, valamint a 0194/2000 sz. FKFP-pályázat.



az összehasonlítandó  $h$  számú csoportban (populációban, statisztikai sokaságban stb.). Ez a nullhipotézis közismerten az egyszempontos varianciaanalízis (VA) módszerével tesztelhető, ha teljesülnek a VA alkalmazási feltételei. (Vincze; 1968. 164. old.)

A társadalomtudományi kutatásokban gyakran előfordul, hogy a függő változó érték-skálája csak legfeljebb az ordinalitás – az értékek nagyság szerinti sorba rendezhetősége – kritériumnak tesz eleget. Ilyen változóhoz jutunk például, ha arra kérjük a vizsgált személyt, hogy ítélje meg egy ötfokú skálán, mennyire kedvel valamely politikai pártot (bizonyos terméket, házastársát, főnökét stb.) a következő fokozatokkal: 1 = nagyon ellenszenves, 2 = enyhén ellenszenves, 3 = közömbös, 4 = enyhén rokonszenves, 5 = nagyon rokonszenves. Hasonlóképpen ordinális skálájúak például az Achenbach-féle Gyermekviselkedési kérdőív (CBCL) háromfokozatú tüneti tételei (1 = soha nem fordul elő, 2 = esetenként előfordul, 3 = gyakran előfordul) és az ezekhez hasonló változók. Ilyen esetekben az átlag használatának jogossága az adatok nagyság szintjének jellemzésére vitatható.

Alternatív középértékként ordinális skálájú változók esetén szóba jöhet a medián is, ez esetben azonban célszerű a folytonosságot megkövetelni, mert diszkrét változóknál a mediánok egyenlősége nem feltétlenül jelenti a változók értékszintjének azonosságát, mint ezt a következő példa is mutatja:

Valószínűség	$P(1)$	$P(2)$	$P(3)$	$P(4)$	$P(5)$
$X$ eloszlása	0,51	0,20	0,10	0,10	0,09
$Y$ eloszlása	0,90	0,06	0,02	0,01	0,01

Ebben a példában az  $X$  és az  $Y$  változó mediánja megegyezik ( $\text{Med}(X) = \text{Med}(Y) = 1$ ), mégis, az  $X$  változó értékei általában nagyobbak, mint az  $Y$  változóé.

A továbbiakban egy olyan statisztikai modellt ismertetünk, amelyben tetszőleges – akár diszkrét, akár folytonos – skálájú változó esetén értelmesen és gyakorlati problémák esetén jól használható módon definiálható az ugyanakkorosság fogalma mindazon összehasonlításokban, amelyek az egy- és a többszempontos VA-ban előfordulnak. Jelen tanulmány központjában  $h$  számú független minta egyszempontos összehasonlítása áll tetszőleges  $h > 2$  mellett, a modell újszerűsége miatt azonban kitérünk a legegyszerűbb,  $h = 2$  esetre is. Ezután definiáljuk a sztochasztikus homogenitás (SZTH) fogalmát, mely az átlagok egyenlőségének egyik alternatívája. Ezzel összefüggésben bemutatunk több statisztikai eljárást, amelyekkel az SZTH nullhipotézise tetszőleges  $h > 2$  mintaszám esetén tesztelhető, majd empirikusan, számítógépes szimulációs vizsgálattal elemezzük ezen próbák érvényességét és hatékonyságát.

### SZTOCHASZTIKUS EGYENLŐSÉG, SZTOCHASZTIKUS KÜLÖNBÖZŐSÉG

Két sokaságot (jelölje őket mondjuk az egyszerűség kedvéért  $\text{Pop}_1$  és  $\text{Pop}_2$ ) valamely  $X$  változó tekintetében nemcsak közvetve, valamely középérték (átlag, medián, trimmelt átlag stb.) által hasonlíthatunk össze, hanem elemeik közvetlen összevetésével is. Például a férfiak és a nők a testmagasságuk szerint úgy is összehasonlíthatók, hogy az összes lehetséges módon összepárosítunk egy férfit és egy nőt, majd meghatározzuk, hogy ezen

összehasonlítások hányad részében magasabb a férfi ( $P(\text{férfi} > \text{nő})$ ) és hányad részében a nő ( $P(\text{nő} > \text{férfi})$ ). Ha ez a két arány megegyezik, akkor azt mondjuk, hogy a férfiak és a nők a testmagasság tekintetében sztochasztikusan egyenlők vagy sztochasztikusan ugyanakkorák.  $P(\text{férfi} > \text{nő}) > P(\text{nő} > \text{férfi})$  esetén azt mondjuk, hogy a férfiak sztochasztikusan magasabbak, ellenkező esetben pedig azt, hogy sztochasztikusan alacsonyabbak, mint a nők.

Teljes általánosságban a sztochasztikus egyenlőség (röviden SZTE) definiáló feltétele, hogy a  $\text{Pop}_1$  sokaságból véletlenszerűen kiválasztott  $X_1$  és a  $\text{Pop}_2$  sokaságból ugyancsak véletlenszerűen és az előzőtől függetlenül kiválasztott  $X_2$  véletlen érték tekintetében fennálljon a

$$P(X_1 > X_2) = P(X_1 < X_2) \quad /1/$$

egyenlőség (Vargha–Czigler; 1999, Vargha; 2000a. 10. fejezet). Analóg módon definiálható az SZTE két tetszőleges eloszlás vonatkozásában is. Ha az SZTE nem áll fenn, akkor sztochasztikus egyenlőtlenségről vagy sztochasztikus különbözőségről beszélünk.

A sztochasztikus különbözőség nagyságának mérésére N. Cliff (1996. 124. old.) a

$$\delta_{12} = P(X_1 > X_2) - P(X_1 < X_2) \quad /2/$$

együtthatót, Vargha A. és H. D. Delaney (1998, 2000) pedig az

$$A_{12} = P(X_1 > X_2) + 0,5 \cdot P(X_1 = X_2) \quad /3/$$

mutatót javasolta. A két mutató egyébként egymással teljesen ekvivalens, mert egymásba egyszerű lineáris transzformációval átalakíthatók:

$$A = (\delta + 1)/2 \quad \text{és} \quad \delta = 2A - 1. \quad /4/$$

Az  $A$  mutató elnevezésére Vargha és Delaney (1998) a valószínűségi fölény mutatója (measure of stochastic superiority), Brunner és Munzel (2000) pedig a relatív kezelési hatás (relative treatment effect) elnevezést javasolja. Ez utóbbi szerzőpáros a sztochasztikusan egyenlő/kisebb/nagyobb fogalmakra a tendenciózusan egyenlő/kisebb/nagyobb (tendentiously equal/smaller/greater) elnevezéseket használja.

A  $\delta$  és az  $A$  mutató definíciója alapján nyilvánvaló, hogy az /1/ formulával definiált SZTE ekvivalens a

$$\delta = 0, \quad A = 0,5 \quad /5/$$

azonosságok fennállásával. E mutatók statisztikai becslésével kapcsolatban lásd: Brunner–Munzel; 2000, Cliff; 1996. 136–140. old. és Vargha; 1999.

Az SZTE hipotézisének vizsgálatára (a sztochasztikus egyenlőtlenség ellenhipotézisével szemben) számos statisztikai próba jöhet számításba.

a) A jól ismert rangsorolós csoport-összehasonlító eljárással, a Mann–Whitney-próbával kapcsolatban már régen kimutatták, hogy az összehasonlított két eloszlás egyenlőségével szembeni bármely ellenhipotézisre

vonatkooán a próba akkor és csakis akkor konzisztens – azaz ereje akkor és csakis akkor tart 1-hez az elem-számok növekedésével bármely rögzített  $\alpha$  szignifikanciaszinten –, ha ezen ellenhipotézis fennállása maga után vonja az összehasonlított két eloszlás sztochasztikus különbözöségét (*Mann–Whitney*; 1947, illetve *Noether*; 1967). A Mann–Whitney-próba azonban csak akkor alkalmazható érvényesen az SZTE önálló tesztelésére, ha bizonyos mellékfeltételek (például az összevetendő két eloszlás alakjának azonossága) teljesülnek.

b) Azzal érvelve, hogy a Mann–Whitney-próba nagy minták esetén ekvivalens egy rangszámokon végrehajtott kétmintás *t*-próbbával, a szóráshomogenitás feltételének sérülése esetén a Mann–Whitney-próba robusztus alternatívájaként *D. W. Zimmerman* és *B. D. Zumbo* (1992, 1993) a rangszámokon végrehajtott Welch-próbát javasolta.

c) Elméleti levezetések és számítógépes szimulációs elemzések azonban arra mutatnak, hogy legalább 10-elemű minták esetén az SZTE egy a *Fligner* és *Policello*, valamint *Cliff* nevéhez fűződő eljárás (*Fligner–Policello*; 1981, *Cliff*; 1996) *Vargha* (2000b) által javasolt módosításával tesztelhető a legmegfelelőbbben.

Megjegyezzük, hogy a sztochasztikus különbözöség becslése, illetve az SZTE tesztelése igen egyszerűen végrehajtható a *MiniStat* programcsomag (*Vargha–Czigler*; 1999) segítségével.

### KETTŐNÉL TŐBB SOKASÁG ÖSSZEHAJONLÍTÁSA: A SZTOCHASZTIKUS HOMOGENITÁS FOGALMA

Tételezzük fel, hogy egy legalább ordinális skálájú *X* változó segítségével akarjuk a  $\text{Pop}_1, \text{Pop}_2, \dots, \text{Pop}_h$  sokaságokat összehasonlítani, ahol  $h > 2$  tetszőleges egész szám. Az egyszerűség kedvéért jelöljük az *X* változót  $X_j$ -vel  $\text{Pop}_j$ -ben ( $j = 1, \dots, h$ ). Ha azt kutatjuk, hogy az *X* változó értékei általában ugyanakkorák-e a vizsgált sokaságokban, akkor kézenfekvő ötlet, hogy megvizsgáljuk: van-e a  $\text{Pop}_j$  sokaságok között olyan, amelyikben az *X* változó értékei általában nagyobbak vagy általában kisebbek, mint a többi sokaságban. Ha nincs ilyen felülről vagy alulról kilógó sokaság, akkor azt mondjuk, hogy a vizsgált  $\text{Pop}_j$  sokaságok az *X* változó tekintetében – s egyben az  $X_1, X_2, \dots, X_h$  változók – sztochasztikusan homogén együttest alkotnak. Ha a sztochasztikus homogenitás nem áll fenn, akkor sztochasztikus heterogenitásról beszélünk.

Matematikailag pontosan a sztochasztikus homogenitás (röviden SZTH) fogalmát a következőképpen definiáljuk. Válasszunk ki először egy sokaságot ( $\text{Pop}_j$ ), s határozzuk meg ennek valószínűségi fölényét a többi sokasággal szemben a /3/ definíció szerint:

$$A_{ji} = P(X_j > X_i) + 0,5 \cdot P(X_j = X_i) \quad (i = 1, \dots, h). \quad /6/$$

Legyen továbbá  $A_{jj} = 0,5$ , hiszen bármely sokaság saját magával nyilván sztochasztikusan egyenlő. Ekkor a

$$p_j = \frac{1}{h} \sum_{i=1}^h A_{ji} \quad /7/$$

mennyiség, a  $\text{Pop}_j$  sokaság ún. sztochasztikus kezelési hatása (stochastic treatment effect) azt mutatja meg, hogy ez a sokaság a többivel a sztochasztikus összehasonlítások során általában nagyobb vagy kisebb-e.  $p_j = 0,5$  esetén a  $\text{Pop}_j$ -beli adatok általában se nem nagyobbak, se nem kisebbek, mint a többiek.

Emiatt az SZTE általánosításaként kettőnél több sokaság sztochasztikus homogenitását a

$$p_1 = \dots = p_h = 0,5 \quad /8/$$

azonossággal definiáljuk.

Ezek a  $p_j$  hatások megadhatók a következőképpen is. A  $h$  számú sokaság közül válasszunk ki véletlenszerűen egyet (mindegyiknek  $1/h$  esélyt adva), majd ebből válasszunk ki véletlenszerűen egy  $X$  megfigyelést. Ezután e választástól függetlenül válasszunk ki a  $\text{Pop}_j$  sokaságból is véletlenszerűen egy  $X_j$  megfigyelést. Ekkor igaz, hogy

$$p_j = A(X_j, X) = P(X_j > X) + 0,5 \cdot P(X_j = X) \quad (j = 1, \dots, h). \quad /9/$$

Bizonyos esetekben célszerű feltételezni, hogy az  $X$  véletlen megfigyelést a  $\text{Pop}_j$  sokaságok összevont együtteséből választjuk ki, melynek során természetesen figyelembe kell venni a sokaságok egymáshoz viszonyított arányát. Ilyen esetben célszerű az SZTH-t valamivel általánosabb formában definiálni. Ehhez csak a sztochasztikus kezelési hatásokat kell kicsit általánosabban meghatározni, a következőképpen:

$$p_j = \sum_{i=1}^h w_i A_{ji} \quad (j = 1, 2, \dots, h), \quad /10/$$

ahol  $w_i$  a  $\text{Pop}_i$  sokaság súlya (aránya). Természetesen a  $w_i$  súlyok összege mindig 1, és a /7/-tel definiált azonos súlyozás esetén  $w_i = 1/h$ .

A /10/ formulából kiolvasható, hogy ha a  $\text{Pop}_j$  sokaságok páronként sztochasztikusan egyenlők (azaz ha  $A_{ij} = 0,5$  minden  $(i, j)$  párra), akkor az SZTH is fennáll, mégpedig bármilyen  $\{w_j\}$  súlyozás mellett. Ennek az állításnak a megfordítása azonban nem igaz, aminek az az oka, hogy a „sztochasztikusan egyenlő” és a „sztochasztikusan kisebb” reláció nem tranzitív. Előfordulhat például olyan ún. „körbeverési” szituáció, amikor a  $\text{Pop}_1$  sokaság sztochasztikusan kisebb a  $\text{Pop}_2$  sokaságnál, a  $\text{Pop}_2$  sokaság sztochasztikusan kisebb a  $\text{Pop}_3$  sokaságnál,  $\text{Pop}_3$  pedig sztochasztikusan nem nagyobb, hanem kisebb  $\text{Pop}_1$ -nél (lásd például Vargha; 2000a. 394. old.).  $\delta$  és  $A$  vonatkozásában ez például azt jelenti, hogy  $\delta_{12} > 0$  és  $\delta_{23} > 0$  nem vonja maga után automatikusan  $\delta_{13} > 0$ , illetve  $A_{12} > 0,5$  és  $A_{23} > 0,5$ , az  $A_{13} > 0,5$  egyenlőtlenség fennállását. Ha azonban az  $X_j$  változók mind szimmetrikus eloszlásúak, akkor az SZTE és a sztochasztikusan kisebb/nagyobb relációk tranzitivitása biztosan teljesül. Ez esetben ugyanis az SZTH ekvivalens a páronkénti SZTE-vel, sőt, az elméleti mediánok egyenlőségével, valamint – ha az átlag értelmes mutató – az elméleti átlagok egyenlőségével is.

Végül megjegyezzük, hogy  $h = 2$  esetén az SZTE és az SZTH ekvivalens relációk, vagyis az SZTH az SZTE egyik – bár nem egyetlen lehetséges – általánosítása.

#### AZ SZTH TESZTELÉSE

Az SZTE vizsgálatához hasonlóan az SZTH tesztelésével kapcsolatban is számos statisztikai eljárás jöhet szóba.

### A Kruskal–Wallis-próba

Matematikailag kimutatták, hogy az összehajonlított  $h$  számú eloszlás egyenlőségével szembeni bármely ellenhipotézisre vonatkozóan a jól ismert Kruskal–Wallis-próba akkor és csakis akkor konzisztens, ha az ellenhipotézis fennállása maga után vonja a vizsgált  $h$  számú eloszlás sztochasztikus heterogenitását, vagyis az SZTH sérülését (Kruskal; 1952, Noether; 1967. 51–52. old.). Ezen összefüggés azonban csak az SZTH egy bizonyos súlyozása esetén érvényes, nevezetesen akkor, ha a /10/ formulában a  $w_i$  súlyokat a következőképpen definiáljuk:

$$w_i = n_i/n. \quad /11/$$

E formulában  $n_i$  az  $i$ -edik független minta elemszáma,  $n = \sum n_i$  pedig az összelemszám.

A Kruskal–Wallis-próba tehát akkor alkalmas az SZTH nullhipotézisének tesztelésére, ha teljesül a következő két feltétel.

1. Az eloszlások az SZTH-n kívül más tekintetben (szóródás, ferdeség, csúcsosság stb.) nem különböznek. Ez a helyzet áll fenn például abban az esetben, ha az eloszlások közötti egyetlen lehetséges eltérés egy esetleges szintkülönbség.

2. A mintaelemszámok – legalább aszimptotikusan – arányosak az összehajonlított sokaságok nagyságával. Ha ismert az egyes sokaságok összsokaságon belüli aránya, akkor ez a feltétel a mintaelemszámok megfelelő megválasztásával biztosítható. Ha viszont ezek az arányok nem ismeretesek, akkor az egyetlen megoldás az, hogy az összsokaságból egyetlen nagy véletlen mintát választunk ki. Ha az összelemszám elég nagy, akkor a mintaelemszámok arányai legalább közelítőleg tükrözni fogják az általuk képviselt sokaságok valódi arányát.

Ha a Kruskal–Wallis-próbát az alkalmazási feltételek bármilyen ellenőrzése nélkül alkalmazzuk, akkor szignifikáns esetben csak egyetlen dolgot állíthatunk: azt, hogy a vizsgált  $X$  változó az összehajonlított sokaságokban nagy valószínűséggel nem ugyanolyan eloszlású.

### Rangszámokon végrehajtott robusztus varianciaanalízisek

Válasszunk ki a  $\text{Pop}_j$  ( $j = 1, 2, \dots, h$ ) sokaságokból egymástól függetlenül rendre  $n_j$  nagyságú véletlen mintákat az  $X$  változóra vonatkozóan. Rangsoroljuk a kapott adatokat a Kruskal–Wallis-próbában előírt módon (Vargha; 2000a. 384–385. old.),  $s$  az egyes minták szerinti csoportok rangszámainak átlagát jelölje rendre  $\bar{R}_1, \bar{R}_2, \dots, \bar{R}_h$ . Matematikailag igazolható, hogy a vizsgált a  $\text{Pop}_j$  sokaságok a /11/ formulában szereplő  $w_j$  súlyokkal akkor és csakis akkor képeznek sztochasztikusan homogén együttest az  $X$  változó tekintetében, ha ezen rangszámátlagok elméleti átlagai (várható értékei) megegyeznek, vagyis ha az összehajonlítandó véletlen minták rangtranszformáltjainak elméleti átlagai ugyanakkorák (a bizonyítást illetően lásd Vargha–Delaney; 1998). Ebből arra következtethetünk, hogy az SZTH tesztelésére szóba jöhet a rangszámokon végrehajtott VA is, ha annak alkalmazási feltételei teljesülnek. Jelen esetben az egyszempontos független mintás VA látszik megfelelőnek, melynek alkalmazási feltételei a következők (Vincze; 1968. 164. old.): a) a minták függetlensége; b) a függő változó normalitása; c) a szóráshomogenitás.

E követelményeket a rangszámadatakra vonatkoztatva, a következőket állapíthatjuk meg.

1. Tekintve, hogy az  $r_{ij}$  ( $j = 1, \dots, h, i = 1, \dots, n_j$ ) rangszámok összege  $n$  elemű össz minta esetén mindig az  $n$ -től függő állandó ( $\sum_j r_{ij} = n(n+1)/2$ ), a rangszáminták sosem függetlenek egymástól, de a rangszámok közötti  $\rho(r_{ij}, r_{kl})$  páronkénti korrelációk az  $n_j$  mintaelemszámok növekedésével 0-hoz tartanak.

2. A rangszámváltozók eloszlása sosem normális, de a VA képleteiben szereplő négyzetes összeg statisztikák bizonyos feltételek esetén aszimptotikusan normális eloszlásúak (Kruskal; 1952, Brunner–Puri; 2001).

3. A rangszámok megőrzik a nagyságrendi viszonyokat, így a nagyobb szóródású minták rangszámai is általában nagyobb szóródásúak lesznek. Ha tehát a szóráshomogenitás az eredeti adatokra teljesül, akkor ez a feltétel többnyire a rangszámok esetén sem sérül számottevően. Ha viszont a szóráshomogenitás feltétele már az eredeti adatokra sem teljesül, akkor ez általában a rangszámokra is átöröklődik.

Mіндеzek alapján az SZTH nullhipotézisének vizsgálatára (a sztochasztikus heterogenitás ellenhipotézisével szemben) közelítő érvényességgel javasolhatók a következő rangsorolós eljárások:

- a szóráshomogenitás feltételének teljesülése esetén a hagyományos egyszempontos VA végrehajtása a rangszámokon (rang VA, röviden rVA), mely egyébként aszimptotikusan ekvivalens a Kruskal–Wallis-próbával (Conover–Iman; 1981);

- a szóráshomogenitás teljesülésétől függetlenül bármely robusztus VA végrehajtása a rangszámokon (robusztus rang VA-k); a robusztus VA-k közül legismertebb a Brown–Forsythe-, a Welch- és a James-féle módszer (Brown–Forsythe; 1974, Welch; 1951, James; 1951).

Hangsúlyoznunk kell, hogy ezek az eljárások heurisztikus alapelven nyugvó közelítő módszerek, amelyek érvényességének körülményeit még empirikusan igazolni kell.

#### *Aszimptotikusan egzakt rangsorolós eljárások*

A nemparaméteres összehasonlításokban nagy jelentőségűek az 1990-es évek közepétől megjelenő és leginkább *M. G. Akritas*, *S. F. Arnold*, *E. Brunner* és *M. L. Puri* nevével fémjelzett tanulmányok. Ezek pontosították a nemparaméteres összehasonlítások egzakt statisztikai nullhipotéziseit és felvázoltak egy olyan matematikai statisztikai modellt, amelyben az ilyen jellegű nullhipotézisek aszimptotikusan – azaz elegendően nagy minták választása esetén – érvényesen tesztelhetők (Akritas–Arnold; 1994, Akritas–Arnold–Brunner; 1997, Brunner–Munzel; 2000, Brunner–Puri; 2001).

Az előbbieken kifejtett SZTH nullhipotézisének vizsgálatára ebben a matematikai modellben E. Brunner egyik tanítványa, *B. Kulle* (1999) dolgozott ki doktori disszertációjában aszimptotikusan érvényes eljárásokat. (Ezek közül a kis mintákra jobb közelítést biztosító VA-típusú módszert a Függelékben részletesen ismertetjük.)

### SZÁMÍTÓGÉPES SZIMULÁCIÓS ELEMZÉS AZ SZTH STATISZTIKAI PRÓBÁIVAL

Az előzőekben több statisztikai próbát tekintettünk át, amelyek szóba jöhetnek az SZTH nullhipotézisének vizsgálatára. Kérdés, hogy melyiket célszerű leginkább alkalmazni, ha valamely empirikus vizsgálatban 3-5 nem túl nagy független mintát szeretnénk egy ordinális skálájú változó nagyság szintjére tekintetében összehasonlítani. A választás

nem könnyű, mert kis minták esetén az áttekintett próbák egyikének érvényességét sem lehet biztosítani, hacsak nem kötünk ki olyan szigorú mellékfeltételeket (például a normalitást vagy azt, hogy az összehasonlítandó eloszlások legfeljebb egy eltolási állandóban különbözhetnek egymástól), amelyek kis minták esetén megbízhatóan nem ellenőrizhetők, és a gyakorlatban legtöbbször nem is teljesülnek (Micceri; 1989). Viszonylag kis – legfeljebb 30 elemű – minták pedig gyakran előfordulnak olyan pszichológiai, biológiai, orvosi, pedagógiai stb. kutatásokban, amelyeknél a megfigyelések költségigénye relatíve magas.

Ilyen esetben csak empirikus úton vethetjük össze a szóba jöhető statisztikai eljárások előnyeit és hátrányait. A matematikai statisztika legfőbb empirikus vizsgáló eszköze a számítógépes szimuláció. Ennek lényege az, hogy a függő változó eloszlására, a szignifikanciaszintre és az elemszámokra vonatkozó rögzített feltételek mellett számítógépes programmal véletlenszerű mintákat alkotunk igen nagy számban ( $I$ ). Ezekben a mintákban végrehajtjuk az elemzésbe bevont statisztikai próbát, döntést hozva – előre rögzített szignifikanciaszinten – a vizsgált nullhipotézis elutasításáról, illetve megtartásáról. Végül meghatározzuk, hogy az  $I$  számú független mintavétel hányad részében utasítottuk el a próba nullhipotézisét. Ha az eloszlást (többmintás összehasonlítások esetén eloszlásokat) úgy állítjuk be, hogy a nullhipotézis teljesüljön, akkor a kapott elutasítási arány a próba elsőfajú hibáját ( $\alpha$ ), ellenkező esetben pedig a próba erejét ( $1 - \beta$ ) becsüli. Ily módon a számítógépes szimulációval lehetővé válik a különféle statisztikai próbák érvényességének és hatékonyságának összehasonlító elemzése.

#### *A szimulációban szisztematikusan változtatott tényezők*

1. *A minták száma.* Jelen szimulációs vizsgálat elsődleges célja olyan általános érvényű összefüggések feltárása volt az SZTH vizsgálatára szóba jöhető különböző statisztikai próbákkal (lásd a 7. pontot) kapcsolatban, amelyek a gyakorlatban előforduló tipikus mintaszámértékek ( $h = 3, 4, 5, \dots$ ) mindegyikénél fellépnek. Ez okból a szimulációk nagy részét a  $h = 3$  mintaszámmal végeztem, majd az egyes következtetések megerősítésére a megfelelő szimulációs elemzést a  $h = 4, h = 5$  mintaszámokkal is végrehajtottam.

2. *A mintaelemszámok.* A szimulációs elemzésekben viszonylag kicsi, 10 és 30 közötti mintaelemszámokat választottam, mert a különféle közelítő eljárások érvényessége többnyire 30 alatti elemszámok esetén a legkritikusabb. Az elemzésekben szerepeltek azonos mintaelemszámok (10-10-10, 20-20-20, 30-30-30), valamint növekvő (10-15-20, 10-20-30) és csökkenő (20-15-10, 30-20-10) elrendezések. Ez lehetővé tette az elemszámok és szórások különböző viszonyainak (például monoton növekvő és monoton fogyó) vizsgálatát is.

3. *A szórás heterogenitás mértéke.* Az elméleti szórások különbözőségének alapvetően 3 szintjét vizsgáltam:

- azonos szórások ( $\sigma_1 = \sigma_2 = \sigma_3$ ),
- közepes mértékű szóráskülönbség,
- nagymértékű szóráskülönbség.

E cél megvalósításához olyan  $\sigma_1:\sigma_2:\sigma_3$  elméleti szórásarányokat választottam, hogy a  $\sigma_1, \sigma_2, \sigma_3$  szórásértékek  $V_e$  variációs együtthatója (lásd Vargha; 2000a. 69–70. old.) azaz

relatív szórása rendre 0, 0,20, illetve 0,40 legyen.<sup>1</sup> Egyes speciális elrendezésekben szélsőségesen nagy,  $Ve = 0,60$  szórás heterogenitást is alkalmaztam. A  $Ve$  variációs együttható tehát itt a szórások szóródását méri. *J. J. Clinch* és *H. J. Keselman* (1982) tanácsát követve ezen szórás heterogenitási szinteket rendre az alábbi  $\sigma_i$  értékekkel értem el:

$Ve$	$\sigma_1$	$\sigma_2$	$\sigma_3$	$\sigma_3/\sigma_1$
0	1	1	1	1
0,20	0,755	1	1,245	1,65
0,40	0,510	1	1,490	2,92
0,60	0,265	1	1,735	6,55

4. *A függő változók eloszlása.* A szimulációban az  $X_1, X_2, X_3$  változók eloszlását külön-külön állítottam be, s itt most csak a folytonos esetre szorítottam.<sup>2</sup> A folytonos eloszlások közül az elemzésekben három szimmetrikus és két ferde eloszlástípust vontam be. A szimmetrikusak képviselőiben szerepelt a normális, a normálisnál lényegesen kisebb csúcsosságú egyenletes, valamint a normálisnál határozottan nagyobb csúcsosságú  $f=5$  szabadságfokú  $t$ -eloszlás,  $t(5)$ . A ferde eloszlásokat egy mérsékelten aszimmetrikus és a normálisnál kismértékben csúcsosabb lambda eloszlás (a lambda eloszlás családdal kapcsolatban lásd *Ramberg–Tadikamalla–Dudewicz–Mykytka*, 1979), valamint az erősen aszimmetrikus és szélsőségesen csúcsos  $f=2$  szabadságfokú  $\chi^2$ -eloszlás képviselte. Mindezen eloszlások ferdeségi és csúcsossági együtthatóját az 1. táblában foglaltuk össze.

1. tábla

<i>A szimulációba bevont folytonos eloszlások ferdeségi és csúcsossági együtthatói</i>		
Eloszlástípus	Ferdeségi együttható $E((X-\mu)^3)/\sigma^3$	Csúcsossági együttható $E((X-\mu)^4)/\sigma^4$
Folytonos egyenletes	0	1,8
Normális	0	3
$t(5)$	0	9
Lambda	0,9	4,2
$\chi^2(2)$	2	9

A szimuláció során először minden eloszlást standardizált (0 átlagú és 1 szórású) alakban állítottam elő, majd ezeket a 3. pontban részletezett szórásokkal besorozva kaptam meg a kívánt szórású változóalakokat.

Az öt eloszlástípust  $h=3$  esetén az alábbi kombinációkban alkalmaztam.

a) *normális*: nnn;

b) *folytonos szimmetrikus*: az egyenletes ( $e$ ), a normális ( $n$ ) és a  $t(5)$ -eloszlás ( $t$ ) következő kombinációi: eee, tt, eet, enn, ent, nne, nnt, tne, tnn, tte;

c) *folytonos aszimmetrikus*: lambda-lambda-lambda és  $\chi^2(2)-\chi^2(2)-\chi^2(2)$ .

<sup>1</sup>  $Ve = \left( \sqrt{\sum (\sigma_i - \bar{\sigma})^2 / N} \right) / \bar{\sigma}$ .

<sup>2</sup> A diszkrét változókkal kapcsolatos elemzésekkel külön tanulmányban foglalkozom.



5. *Az SZTH versus sztochasztikus heterogenitás.* Szimmetrikus eloszlások esetén az SZTH ekvivalens az elméleti átlagok egyenlőségével. Ezért ebben az esetben az SZTH fennállását azonos (mellesleg 0-val egyenlő) elméleti átlagok beállításával értem el minden szórás-kombinációban. Azonos ferde eloszlások esetén viszont az SZTH csak akkor teljesül, ha az összehasonlított eloszlások átlaga és szórása egyaránt azonos. Ezért ez esetben azonos átlagok és  $Ve > 0$  esetén az SZTH bizonyosan sérül. Emiatt az SZTH fennállását különböző szórású ferde eloszlások esetén (azonos súlyok alkalmazásával) alkalmas eltolási állandók felhasználásával értem el, amelyeket egy iterációs algoritmusú számítógépes programmal empirikusan határoztam meg.

2. tábla

*Az SZTH fennállását azonos súlyokkal biztosító eltolási állandók az 1. táblában szereplő lambda és  $\chi^2$  eloszlás esetén különböző szórás-heterogenitási szinteken*

$Ve$	Lambda eloszlás			$\chi^2(2)$ eloszlás		
	1. minta	2. minta	3. minta	1. minta	2. minta	3. minta
0,20	-0,1120	-0,0575	0	-0,2220	-0,1120	0
0,40	-0,2115	-0,1105	0	-0,4030	-0,2067	0
0,60	-0,2880	-0,1500	0	-0,5450	-0,2880	0

A szimulációban a szignifikáns eredmények aránya az SZTH-t beállítva a próba első-fajú hibájára ad pontbecslést, míg sztochasztikus heterogenitást beállítva a próba erejéről kapunk tájékoztatást. A sztochasztikus heterogenitást úgy értem el, hogy az SZTH-t biztosító  $X_1, X_2, X_3$  változókban az első változóból kivontam 0,40-ot, a harmadikhoz pedig hozzáadtam ugyanennyit. Különböző elemszámok esetén megvizsgáltam az ellenkező előjelű  $\mu_1 = 0,40, \mu_2 = 0, \mu_3 = -0,40$  beállítás hatását is. Ezen esetekben azonos szórások (azaz  $Ve = 0$ ) esetén az 1. és a 3. minta elméleti eloszlása közötti különbség szokásos Cohen-féle  $\Delta$  hatásmértéke éppen 0,8, amely a Cohen-féle konvenció szerint igen jelentősnek tekinthető (Cohen; 1977. 20. old.). Az SZTH szempontjából különösen érdekes ferde eloszlások esetén megvizsgáltam a  $\Delta = 0,5$  esetet is, mely *J. Cohen* szerint a közepes szintű eltérés mértéke, s amelyet az  $X_1 - 0,40, X_3 + 0,1$  eltolásokkal értem el.

6. *Szignifikanciaszint.* A szimulációs elemzéseket a gyakorlatban legtöbbször előforduló  $\alpha = 0,01, \alpha = 0,05, \alpha = 0,10$  szintek mindegyikére elvégeztem.

7. *Statistikai próbák.* A szimulációs elemzésekbe a következő próbákat vontam be.

i) *Egyszempontos független mintás VA.* Ennek érvényessége elméletileg csak az nnn eloszláskombinációra garantált  $\sigma_1 = \sigma_2 = \sigma_3$ , azaz  $Ve = 0$  esetén. Más szimmetrikus eloszláskombinációra  $Ve = 0$  esetén a VA legfeljebb csak közelítő pontosságú. Ferde eloszlások esetén az SZTH fennállását csak az biztosítja, ha az  $X_1, X_2, X_3$  változók azonos eloszlástípusba tartoznak, és  $Ve = 0$ , így a VA is csak ez esetben jöhet szóba az SZTH tesztelésére.

ii) *Egyszempontos független mintás robusztus VA-k: Brown-Forsythe-próba (BF), Welch-próba (W), James-próba (J).* E próbák érvényessége a normalitás esetén és nagy mintákra garantált. Más szimmetrikus eloszláskombinációk esetén csak legfeljebb közelítő pontosságra számíthatunk. Azonos ferde eloszlások esetén az SZTH tesztelése e próbák kapcsán is csak a  $Ve = 0$  feltétel teljesülése esetén jöhet szóba. (Végrehajtásuk tech

nikai részleteire terjedelmi okokból nem térünk ki. A részleteket lásd a BF-fel kapcsolatban: *Brown–Forsythe*; 1974, *Dixon*; 1990. 547. old. vagy *Vargha*; 1981. 94–95. old., a W-vel kapcsolatban: *Welch*; 1951 vagy *Wilcox*; 1996. 183–184. old., a J-vel kapcsolatban: *James*; 1951 vagy *Wilcox*; 1996. 182–183. old.).

Ezeket a robusztus próbákat a hagyományos VA alternatíváiként normális eloszlású változókra dolgozták ki az elméleti átlagok egyenlőségének tesztelésére a szóráshomogenitás feltételének sérülése esetén. Emiatt normális eloszlású változók esetén segítségükkel az átlagok egyenlőségével ez esetben ekvivalens SZTH is tesztelhető. Más szimmetrikus eloszláskombinációkra e próbák legfeljebb csak közelítő pontosságúak. Ferde eloszlások esetén az SZTH fennállását csak az garantálja, ha az  $X_1, X_2, X_3$  változók azonos eloszlástípusba tartoznak, és  $V_e = 0$ , így csak ez esetben jöhetnek szóba az SZTH tesztelésére.

iii) *Kruskal–Wallis-próba (KW)*. Végrehajtásának lépéseit illetően lásd *Kruskal–Wallis* (1952) vagy *Vargha* (2000a. 384–386. old.).

iv) *Rangszámokon végrehajtott egyszempontos független mintás VA (rang VA)*. A rang VA (röviden rVA) végrehajtásának módja:

- rangszámok képzése ugyanúgy, mint KW-ben,
- a rangszámokon egyszempontos független mintás VA végrehajtása.

v) *Rangszámokon végrehajtott robusztus egyszempontos független mintás VA-k (rBF, rW, rJ)*. E próbák végrehajtásának módja:

- rangszámok képzése ugyanúgy, mint KW-ben,
- a rangszámokon végrehajtuk BF-et, W-t, illetve J-t.

vi) *Kulle-féle próbák*. Mintaelemszámmal súlyozott sokaságok/eloszlások esetén az SZTH tesztelése a

$$K_s = \hat{f} \cdot F_K \quad /12/$$

próbat statisztika segítségével hajtandó végre (lásd /F4/ és /F5/ formula), mely az SZTH fennállása esetén aszimptotikusan  $\hat{f}$  szabadságfokú  $\chi^2$  eloszlást követ. Ezt a próbát a továbbiakban KKF-fel rövidítjük. Azonos súlyok esetén az SZTH tesztelése a

$$K = \hat{g} \cdot G_K \quad /13/$$

próbat statisztika segítségével hajtandó végre (lásd /F19/ és /F20/ formula), mely az SZTH fennállása esetén aszimptotikusan  $\hat{g}$  szabadságfokú  $\chi^2$  eloszlást követ. Megjegyzendő, hogy azonos mintaelemszámok esetén a két próbat statisztika, valamint a két szabadságfok megegyezik, így a két eljárás ekvivalens. Ezt a próbát a továbbiakban KKG-vel rövidítjük.

A normális eloszlású változókkal végzett szimulációs vizsgálatban *B. Kulle* (1999. 70. old.) arra az eredményre jutott, hogy előbbi próbái 30-nál kisebb átlagos mintaelemszám esetén liberálisak (azaz elsőfajú hibájuk többnyire számottevően magasabb, mint a szignifikanciaszint). Ez a jelenség gyakori az aszimptotikusan normális eloszlású próba

statisztikák esetén, amelyek kis minták mellett pontosabban közelíthetők egy alkalmas szabadságfokú  $t$ -eloszlással (Brunner–Munzel; 2000, Vargha; 2000b, Vargha–Delaney; 2000). Emiatt a Kulle által javasolt próbák mellett a következő módosított változatokat is bevontam a szimulációs elemzésekbe.

a)  $K_s$  helyett az /F4/ formulával definiált  $F_K$  próbastatisztikát vettem, s ezt az  $F$ -táblázat szerint értékeltem ki az  $f_1 = \hat{f}$ ,  $f_2 = f_w$  szabadságfokokkal, ahol  $\hat{f}$ -ot az /F5/,  $f_w$ -t pedig az alábbi formula segítségével határoztam meg:

$$f_w = \left[ \frac{3}{h^2 - 1} \sum_{j=1}^h \frac{(1 - c_j / U)^2}{n_j - 1} \right]^{-1}; \quad /14/$$

itt  $c_j = n_j / \hat{\sigma}_j^2$ , ahol  $\hat{\sigma}_j^2$  az /F11/ képlet segítségével definiált mennyiség és  $U = \Sigma c_j$ . E próbát a továbbiakban KF1-gyel jelöljük. Ezen eljárás a következő megfontoláson alapul. Kulle (1999. 36–37. old.) matematikai levezetéssel igazolta, hogy /12/ próbastatisztikájú próbája  $h = 2$  esetén az SZTE tesztelésére Brunner, Puri és Sun (1995) által javasolt próbára redukálódik, mely egyébként ugyanaz, mint Brunner és Munzel (2000) nagy mintákra érvényes eljárása. Ez utóbbiak szimulációs vizsgálattal kimutatták, hogy ennek a kétmintás próbának az érvényessége kis és közepes minták ( $n_j < 50$ ) esetén számottevően javítható, ha próbastatisztikáját nem a standard normális, hanem a  $t$ -eloszlás táblázatával értékeljük ki. Az alkalmazandó szabadságfokot a kétmintás Welch-próba szabadságfokára vonatkozó képlettel határozták meg oly módon, hogy az eredeti adatokból számított mintavariációk helyett az /F3/ formulával definiált  $Z_j$  változók /F11/ formulában szereplő  $\hat{\sigma}_j^2$  varianciabecsléseit vették (Brunner–Munzel; 2000. (5.9) formula). Ennek analógiájára a /14/ formula a Welch-féle VA próbastatisztikájának  $f_2$  szabadságfokát definiáló képlet azzal az eltéréssel, hogy benne az eredeti adatok mintavariációi helyett szintén a  $\hat{\sigma}_j^2$  varianciabecslések szerepelnek.

b) Hasonló megfontolások alapján a /13/-ban szereplő  $K$  statisztika helyett az /F19/ formulával definiált  $G_K$  próbastatisztikát vettem, s ezt az  $F$ -táblázat szerint értékeltem ki az  $f_1 = \hat{g}$ ,  $f_2 = g_w$  szabadságfokokkal;  $\hat{g}$ -ot az /F20/,  $g_w$ -t pedig a /14/ formula segítségével határoztam meg a  $c_j = n_j / \hat{\sigma}_j^2$  behelyettesítéssel, ahol  $\hat{\sigma}_j^2$  az /F26/ képlete segítségével definiált mennyiség és  $U = \Sigma c_j$ . E próbát a továbbiakban KG1-gyel jelöljük.

Néhány próbaszimulációt elvégezve úgy tűnt, hogy  $F_K$  és  $G_K$  esetenként a névlegesnél kissé alacsonyabb elsőfajú hibákat eredményez, ezért a szimulációban kipróbáltam két olyan változatot is, amelyben az  $f_w$ , illetve a  $g_w$  szabadságfokot megnöveltem, a következőképpen:

$$f_{w2} = \min(2f_w, n - h) \text{ és } g_{w2} = \min(2g_w, n - h). \quad /15/$$

Ezen újabb szabadságfok-változatok értéke tehát általában a korábbiak kétszerese, de felülről limitáljuk őket a hagyományos VA  $F$  próbastatisztikájának  $f_2 = n - h$  szabadságfokával. Az így kapott próbaváltozatokat a továbbiakban KF2-vel, illetve KG2-vel jelöljük.

### *A szimulációs elemzés néhány technikai részlete*

A szimulációs elemzések programját Turbo Pascal 7.0 programozási nyelven írtam, és a futtatásokat Pentium PC-n végeztem el. E programmal az előbb leírt elemszám-, variancia- és eloszláselrendezések mindegyikében „véletlen” megfigyeléseket iktattam be, egy-egy elrendezésben 10 000 egymástól független mintavétellel. Ez az ismétlési szám az elsőfajú hiba, illetve a próba erejének legalább 0,0050 pontosságú becslését tette lehetővé (a becslés standard hibája egészen pontosan 10 százalékos valódi elsőfajú hiba mellett 0,0030, 5 százalékos elsőfajú hiba mellett 0,0022, 1 százalékos elsőfajú hiba mellett pedig 0,0010). Minden mintavétel esetén kiszámítottam az előbbieken részletezett 15 statisztikai próba mindegyikét, és megnéztem, hogy szignifikánsak-e 1, 5, illetve 10 százalékos szinten, majd meghatároztam a szignifikáns eredmények arányát a 10 000 mintavételben mindhárom valószínűségi szintre vonatkozóan. Ezek olyan eloszláselrendezésekben, amelyekben az SZTH fennáll, a vizsgált próbák elsőfajú hibáját, egyébként az erejét becsülik.

Az egyenletes eloszlású pszeudovéletlen megfigyeléseket a Turbo Pascalba beépített Random függvénnyel hoztam létre. Ez lineáris kongruens véletlen szám generátor, mely egy közelmúltban megjelent tanulmány szerint több vonatkozásban a legjobb eljárások egyikének tekinthető (Onghena, 1993). A normális, a  $t$ - és a khi-négyzet-eloszlást a *W. H. Press, B. P. Flannery, S. A. Teukolsky és W. T. Vetterling* (1989) könyvében, a szóban forgó lambda eloszlást pedig a *J. S. Ramberg, P. R. Tadikamalla, E. J. Dudewicz és E. F. Mykytka* (1979) tanulmányában leírt módon hoztam létre az említett módszerrel kapott egyenletes eloszlású pszeudovéletlen számok segítségével. A lambda eloszlásokkal részletesebben Vargha A. (1996) írása foglalkozik.

### *A szimulációs elemzések eredményei*

*1. Normális eloszlás azonos elméleti szórásokkal: a szimulációs eljárás hitelessége.* Az elemzések csak akkor lehetnek hitelesek, ha a kapott eredmények nem mondanak ellent elméleti ismereteinknek. Esetünkben ez többek között azt jelenti, hogy a VA feltételeinek (normális eloszlás és szóráshomogenitás) teljesülése esetén a kapott elsőfajú hibabecslések nem különbözhetnek számottevően a szignifikanciaszinttől. E feltételek mellett az elsőfajú hibabecslések a három szinten a következő tartományokba estek:

$$\alpha = 0,01: 0,0079 - 0,0112;$$

$$\alpha = 0,05: 0,047 - 0,054;$$

$$\alpha = 0,10: 0,098 - 0,105.$$

Ezen értékek közül csak az 1 százalékos szinten kapott 0,0079-es minimális érték különbözik szignifikánsan a névleges 0,01-os értéktől (nagy mintás binomiális próbával  $z = 2,10, p < 0,05$ ). A hitelesség ellenőrzésére a szóban forgó elrendezéssel újabb szimulációs elemzést végeztem, de 10 000 helyett 100 000 véletlen mintavétellel. Az ekkor kapott 0,0096-es hibabecslés már nem tért el szignifikánsan az  $\alpha = 0,01$ -os névleges szinttől.

Normális eloszlás és  $\sigma_1 = \sigma_2 = \sigma_3$  esetén a VA mellett a KKF és a KKG kivételével teljesen elfogadható elsőfajú hibabecsléseket kaptunk az összes próbára 5 és 10 százalé

kos szignifikanciaszinten; az elsőfajú hibabecslések  $\alpha = 0,05$  esetén 0,045 és 0,057,  $\alpha = 0,10$  esetén pedig 0,092 és 0,112 közé estek. 1 százalékos szinten a becslések szóródása nagyobb volt (értéktartomány: 0,0071 – 0,0170). A KKF és a KKG Kulle eredményeivel összhangban végig érezhetően liberális, vagyis a névlegesnél magasabb elsőfajú hibákat adott (a kapott elsőfajú hibabecslések tartománya a három szinten rendre 0,014 – 0,023, 0,052 – 0,068, illetve 0,099 – 0,117). Emiatt a továbbiakban eltekintünk a KKF-re és a KKG-re vonatkozó eredmények ismertetésétől.

2. *Folytonos szimmetrikus eloszlások azonos mintaelemszámokkal.* Ha a normalitástól eltekintünk, akkor már nem feltétlenül igaz az, hogy a szimulációba bevont eljárások minden esetben érvényes eljárások lesznek. Ez indokolja leginkább a szimulációs elemzéseket, amelyek segítségével empirikusan győződhetünk meg az egyes eljárások érvényességi feltételeiről. Az érvényesség fennállása erősen függ a szórásheterogenitás mértékéről, ezért az eredményeket e szempont szerint bontva ismertetjük.

a) *Azonos elméleti szórások ( $Ve = 0$ ).* A megvizsgált 11 eloszláskombináció (eee, enn, ent stb.) és a három azonos elemszám-kombináció (10-10-10, 20-20-20, 30-30-30), vagyis összesen 33 elrendezés esetén nyert minimális és maximális elsőfajú hibabecsléseket 5 százalékos névleges szint mellett a 3. tábla tartalmazza három szórásheterogenitási szintre vonatkozóan ( $Ve = 0$ ,  $Ve = 0,2$ ,  $Ve = 0,4$ ).

3. tábla

33 elrendezés esetén nyert minimális és maximális elsőfajú hibabecslések  
 $\alpha = 0,05$  névleges szint mellett három szórásheterogenitási szinten azonos mintaelemszámok esetén

Próba	$Ve = 0$		$Ve = 0,2$		$Ve = 0,4$	
	minimum	maximum	minimum	maximum	minimum	maximum
VA	0,045	0,057	0,045	0,058	0,055	0,071**
BF	0,043	0,057	0,045	0,057	0,049	0,064*
W	0,043	0,057	0,043	0,054	0,044	0,059
J	0,040	0,057	0,038	0,053	0,038	0,053
KW	0,041	0,056	0,043	0,060*	0,053	0,070**
RF	0,046	0,057	0,046	0,062*	0,056	0,073**
RBF	0,045	0,057	0,046	0,061*	0,055	0,070**
RW	0,048	0,058	0,047	0,061*	0,054	0,072**
RJ	0,047	0,058	0,047	0,060*	0,053	0,069**
KF1=KG1	0,039	0,054	0,041	0,052	0,043	0,051
KF2=KG2	0,044	0,056	0,044	0,055	0,047	0,059

\* Szignifikánsan nagyobb 0,055-nél.

\*\* Szignifikánsan nagyobb 0,060-nél.

Megjegyzés. KF1 és KG1, illetve KF2 és KG2 itt most az azonos elemszámok miatt egyezik meg.

A 3. táblából kiolvashatjuk, hogy a kapott elsőfajú hibabecslések tartománya  $Ve = 0$  és  $\alpha = 0,05$  esetén 0,039 – 0,058, ahol a felső határ nem nagyobb szignifikánsan a névleges szint 110 százalékánál (azaz 0,055-nél), mely a szokásosan elfogadott Bradley-féle konvenció szerint a robusztusság szigorú kritériumának tekinthető (Bradley; 1978). Ugyanez a tartomány  $\alpha = 0,1$  esetén 0,091 – 0,112, s a felső határ itt sem nagyobb szignifikánsan a névleges szint 110 százalékánál. Mindez azt jelenti, hogy folytonos és szim

metrikus eloszlások körében  $Ve = 0$  esetén a 3. táblában feltüntetett próbák egyikének sem sérül számottevően az érvényessége a szokásos 5 százalékos szignifikanciaszinten, ha azonos mintaelemszámokkal dolgozunk.  $\alpha = 0,01$  esetén azonban W és a rangsorolási eljárások többsége számos elrendezésben érezhetően liberális, a névleges szintet néha 50-60 százalékosan meghaladó elsőfajú hibát adva. Emiatt a továbbiakban csak az  $\alpha = 0,05$ ,  $\alpha = 0,1$  szintekhez tartozó eredményeket részletezzük.

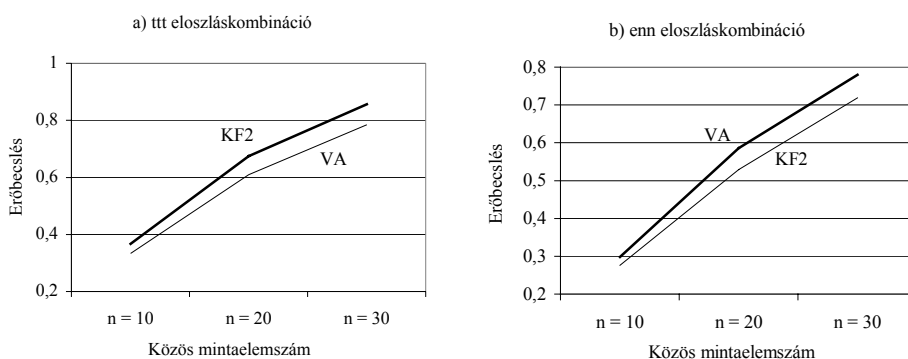
Az egyazon nullhipotézist tesztelő és nagyjából hasonló érvényességű próbák közül nyilván azt érdemes a gyakorlatban használni, amelyiknek ereje a legnagyobb. Az erőelemzés esetünkben a következő eredményekre vezetett. 5 és 10 százalékos szignifikanciaszinten a paraméteres próbák (VA, BF, W és J) közül legnagyobb erejű többnyire a VA, de a többi próba ereje sem számottevően kisebb, az erőkülönbség sosem nagyobb 0,02-nél. A rangsorolási próbák (KW, rVA, rBF, rW, rJ, KF1 és KF2) közül a legerősebb többnyire az rW és a KF2, de a KF1 kivételével a többi ereje sem számottevően kisebb. Érthető okokból a KF1 ereje általában 0,02-0,03-dal kisebb, mint a KF2-é, illetve a többi rangsorolási próbáé. Ez okból a paraméteres és a rangsorolási próbák ereje közötti összevetéshez elegendő a VA és a KF2 erejének összehasonlítása. Ez az elemzés a következő összefüggésekre világított rá.

i) A ttt eloszláskombináció esetén a KF2 minden szignifikanciaszinten számottevően erősebb, mint a VA (az 5 százalékos szinttel kapcsolatban lásd az 1. ábra a) grafikonját). Annak ellenőrzésére, hogy ez a jelenség háromnál több minta összehasonlítása esetén is fennáll-e, azonos  $n = 15$  elemszámmal szimulációt végeztem a  $h = 3, 4, 5$  mintaszámokkal. A kapott eredmények megerősítették a feltételezést. A KF2 és a VA ereje közötti különbség  $\mu_1 = -(0,40 + 0,05 \cdot (h-3))$ ,  $\mu_i = 0$  ( $1 < i < h$ ),  $\mu_h = 0,40 + 0,05 \cdot (h-3)$  mellett a három elemszámszinten rendre 0,06, 0,06, illetve 0,05 volt, KF2 javára.

ii) nnt és tnn esetén VA és KF2 ereje hasonló nagyságrendű, az eltérés sosem nagyobb, mint 0,02.

iii) Az összes többi kombinációban (amelyekben az egyenletes eloszlás – e – legalább egyszer szerepel) a VA számottevően erősebb, mint a KF2. Illusztrációként az enn kombinációra vonatkozó összehasonlítás az 1. ábra b) grafikonján látható. Az 1. pontban említettel azonos elemzés megerősítette, hogy ez az összefüggés  $h = 4$  és  $h = 5$  esetén is fennáll: enn esetében a KF2 és a VA ereje közötti különbség a három elemszámszinten rendre 0,04, 0,06, illetve 0,05 volt, VA javára.

1. ábra. Erőbecslések két eloszláskombinációra  $\alpha = 0,05$  szinten



b) Mérsékeltlen különböző szórások ( $Ve = 0,2$ ). Közepes mértékű szórás heterogenitás esetén (ekkor  $\sigma_{\max}/\sigma_{\min} = 1,65$ ) az erőre vonatkozó eredmények a következőképpen alakultak. A paraméteres próbák közül a W tűnik a legjobbnak, amelynek ereje több esetben

(főleg nagyobb minták esetén) meghaladja a VA erejét. Hasonló szerepváltás figyelhető meg az rW és a KF2 között, bár ez részben abból is fakad, hogy az rW elsőfajú hibaszintje a szórás heterogenitás miatt enyhén megemelkedett. (Lásd a 3. táblát.) A paraméteres-rangsorolásos összevetésben most a következő összefüggések tűntek ki.

i) A ttt eloszláskombinációban ismét a rangsorolásos próbák szerepeltek a legjobban. Például  $\alpha = 0,05$  esetén a KF2 és a W ereje közötti különbség a három elemszám szinten rendre 0,03, 0,04, illetve 0,05, míg ugyanez a különbség az rW és a W között mindhárom elemszám szinten 0,07 lett.

ii) A többi olyan eloszláskombinációban, ahol  $t$  volt a 3. helyen, vagyis a legnagyobb szórással párosítva (nnt, ent, eet), a paraméteres és a rangsorolásos próbák erejének nagysága hasonló volt, a rangsorolásos próbák esetenkénti enyhe fölényével.

iii) Azokban az eloszláskombinációkban, ahol  $e$  volt a legnagyobb szórással párosítva (nne, tne, tte, eee), a paraméteres próbák egyértelmű fölényben voltak a rangsorolásos próbákkal szemben. Például a tte kombinációban  $\alpha = 0,05$  szinten az erőkülönbség a három eloszlásszinten a W-rW viszonylatban rendre 0,02, 0,07 és 0,08, a W-KF2 összevetésben pedig rendre 0,05, 0,12 és 0,13 volt.

c) *Erősen különböző szórások* ( $Ve = 0,4$ ). A 3. tábla utolsó oszlopából kiolvasható, hogy nagymértékű szórás heterogenitás esetén (ekkor  $\sigma_{\max}/\sigma_{\min} = 2,92$ ) érvényesként az SZTH tesztelésére csak W, J, KF1 és KF2 jöhet szóba. Mivel a két paraméteres próba (W és J) közül mindig W, a két Kulle-féle eljárás közül pedig mindig KF2 az erősebb, ezért az SZTH tesztelésére legalkalmasabb próba kiválasztásához ez esetben a W és a KF2 erejét kell összevetni. Ez az összehasonlítás a következő eredményre vezetett.

i) A ttt eloszláskombinációban W és KF2 ereje azonos nagyságú. Például 5 százalékos szignifikanciaszinten W becslt ereje a három elemszám szinten rendre 0,35, 0,64 és 0,82, míg KF2-é 0,33, 0,63, illetve 0,82 volt.

ii) Mérsékelt W erőfölény figyelhető meg azokban az eloszláskombinációkban, ahol a legnagyobb szórás magas csúcossággal jár együtt (nnt, ent, eet stb.).

iii) Igen nagy W erőfölény figyelhető meg az összes többi eloszláskombinációban (nnn, tte, eee stb.). Például tne kombinációban 5 százalékos szignifikanciaszinten W becslt ereje a három elemszám szinten rendre 0,29, 0,59 és 0,79, míg a KF2-é 0,19, 0,35, illetve 0,51. Érdekes, hogy ebben az esetben még a megemelkedett elsőfajú hibájú rW is észlelhetően gyengébb W-nél (rW becslt ereje a három elemszám szinten rendre 0,27, 0,49, illetve 0,67).

Az rW minden helyzetben tapasztalható erőfölénye KF2-vel szemben magas szórás heterogenitás esetén azt az ötletet adta, hogy talán lehet rW-t konzervatív irányban módosítani (vagyis a névleges szintet többnyire meghaladó elsőfajú hibáját mérsékelni) úgy, hogy KF2-vel szembeni erőfölénye megmaradjon. Ez okból egyes szimulációs elemzésekben az rW eljárás  $f_2$  szabadságfokát a következőképpen módosítottam:

$$f_{2uj} = f_2 \cdot (1,1 - Ve(s)), \quad /16/$$

ahol  $Ve(s)$  az összehasonlított független minták tapasztalati szórásainak variációs együtthatója. Emellett  $f_{2uj}$  értéktartományát felülről  $f_2$ -vel, alulról pedig  $0,5 \cdot f_2$ -vel korlátoztam. Minél nagyobb a vizsgált eloszlások szórásai közötti különbség, annál nagyobb általában  $Ve(s)$ , s annál kisebb az  $f_{2uj}$  szabadságfok  $f_2$ -höz viszonyítva. rW ezen módosítását a továbbiakban rW2-vel jelöljük.

Az rW2 alkalmasságának ellenőrzésére a következő elemzést végeztem.  $I = 10\,000$  véletlen mintavétellel az összes szimmetrikus eloszlásra vonatkozó szimulációs elrende

zésből (ezek száma 209 volt) kikerestem azt a 25-öt, amelyre W elsőfajú hibabecslése  $\alpha = 0,05$  szinten a legnagyobb volt. E becslésértékek 0,064 és 0,072 közé estek. Ezzel a 25 elrendezéssel újabb szimulációt végeztem rW-re és rW2-re vonatkozóan, és a következő eredményt kaptam. rW ezen szimulációkban is erősen liberálisnak mutatkozott  $\alpha = 0,05$  szinten 0,060 és 0,073 közé eső elsőfajú hibákkal. Ugyanakkor rW2 elsőfajú hibabecslése a 25-ből csak mindössze két esetben haladta meg a 0,060-et (egy esetben 0,065, egy másikban pedig 0,062 volt). E maradék két esetben újabb, de jóval nagyobb ismétlési számú ( $I = 100\ 000$ ) szimulációval becsültem meg rW2 elsőfajú hibáját, amely most 0,062-nek, illetve 0,059-nek adódott. Mindez azt mutatja, hogy az rW eljárás  $f_2$  szabadságfokának javasolt módosítása elfogadható mértékűre csökkenti az inflációs torzítást. Ugyanakkor rW2-vel is lefuttatva a tt eloszláskombinációban a többi próbával már elvégzett szimulációkat, kiinduló feltételezésemet megerősítő eredményeket kaptam. Ezekben az elrendezésekben ugyanis rW2 határozottabban nagyobb erejű volt KF2-nél és W-nél, s ugyanezt mutatták a  $h > 3$  mintaszámokkal végrehajtott szimulációk is. Például  $n = 15$  és  $\alpha = 0,05$  esetén  $h = 3, 4, 5$  értékek mellett W elsőfajú hibabecslése rendre 0,50, 0,63, 0,70, a KF2-é 0,49, 0,55, 0,58, az rW2-é pedig 0,55, 0,69, 0,75 volt.

3. *Folytonos szimmetrikus eloszlások különböző mintaelemszámokkal.* Mivel 10 és 30 közötti elemszámok segítségével a szóráshomogenitás fennállása, mely feltételre nézve VA, rVA és KW különböző mintaelemszámok esetén fokozottan érzékeny (lásd például Scheffé; 1959. 353. old.), nem ellenőrizhető kellő megbízhatósággal, ez esetben csak a robusztus próbák elemzésére szorítkozunk. Az 5 és 10 százalékos szintre vonatkozó elsőfajú hibabecslések minimumát és maximumát a 4. tábla tartalmazza az elemszámok és az elméleti szórások közötti kapcsolat jellege szerinti bontásban. E táblából kiolvasható, hogy a robusztus rang VA-k a szóráshomogenitás sérülése esetén gyakran idéznek elő a névlegesnél magasabb elsőfajú hibát, s ugyanez igaz a Kulle-féle próbákra, ha az elemszámok és az elméleti szórások negatív kapcsolatban vannak egymással. A paraméteres próbák közül láthatóan W és J tartja legjobban a névleges szintet.

4. tábla

Tíz robusztus próba minimális és maximális elsőfajú hibabecslése különböző elemszámok esetén az elemszámok és elméleti szórások közötti kapcsolat jellege szerinti bontásban

Próba	Az elemszámok és az elméleti szórások közötti kapcsolat					
	negatív ( $k = 44$ )		$V_e = 0$ ( $k = 22$ )		pozitív ( $k = 44$ )	
	minimum.	maximum	minimum	maximum	minimum	maximum
	$\alpha = 0,05$					
BF	0,045	0,066**	0,044	0,053	0,047	0,064*
W	0,042	0,060*	0,045	0,056	0,044	0,056
J	0,038	0,055	0,042	0,052	0,041	0,053
rBF	0,049	0,068**	0,047	0,056	0,049	0,072**
rW	0,051	0,069**	0,050	0,059*	0,052	0,070**
rJ	0,049	0,065**	0,047	0,056	0,050	0,069**
KF1	0,044	0,058	0,045	0,052	0,039	0,047
KF2	0,053	0,068**	0,052	0,059*	0,044	0,055
KG1	0,046	0,062*	0,046	0,053	0,041	0,049
KG2	0,053	0,071**	0,051	0,061*	0,046	0,056

(A tábla folytatása a következő oldalon.)



(Folytatás.)

Próba	Az elemszámok és az elméleti szórások közötti kapcsolat					
	negatív ( $k = 44$ )		$Ve = 0$ ( $k = 22$ )		pozitív ( $k = 44$ )	
	minimum	maximum	minimum	maximum	minimum	maximum
	$\alpha = 0,10$					
BF	0,093	0,113	0,090	0,103	0,096	0,110
W	0,092	0,108	0,090	0,105	0,091	0,105
J	0,088	0,104	0,089	0,104	0,091	0,105
rBF	0,096	0,117*	0,096	0,106	0,099	0,125**
rW	0,098	0,124*	0,098	0,109	0,100	0,129**
rJ	0,096	0,122*	0,097	0,108	0,099	0,129**
KF1	0,089	0,102	0,089	0,098	0,086	0,096
KF2	0,098	0,112	0,097	0,107	0,092	0,105
KG1	0,090	0,105	0,090	0,099	0,086	0,100
KG2	0,098	0,114	0,098	0,108	0,093	0,107

\* Szignifikánsan nagyobb  $1,1\alpha$ -nál.

\*\* Szignifikánsan nagyobb  $1,2\alpha$ -nál.

Az előző, 2. pontban az rW módosított formájával, rW2-vel kapcsolatban említett szimulációs elemzések alapján rW2-t is a névleges szintet nem számottevően meghaladó elsőfajú hibájú eljárások közé kell sorolnunk. Így az SZTH tesztelésére ez esetben legalkalmasabb próba megtalálásához leginkább a W, az rW2 és a Kulle-féle próbák erejét kell összevetnünk különböző elrendezésekben. A legjelentősebb összefüggések a legnagyobb szórás heterogenitási szinten jelentkeztek, melyek szemléltetésére a három eloszláskombinációval kapcsolatos erőelemzés eredményeit az 5. tábla mutatja be.

5. tábla

Négy robusztus próba erőbecslései különböző elemszámok esetén magas szórás heterogenitási szinten ( $Ve = 0,4$ ) három eloszláskombinációra vonatkozóan

$n_1-n_2-n_3$	W	rW2	KF2	KG2
	próba			
	tnn eloszláskombináció			
10-15-20	0,50	0,48	0,43	0,41
10-20-30	0,62	0,60	0,57	0,55
20-15-10	0,39	0,37	0,28	0,29
30-20-10	0,46	0,43	0,29	0,30
	tnu eloszláskombináció			
10-15-20	0,49	0,38	0,32	0,31
10-20-30	0,62	0,49	0,44	0,42
20-15-10	0,37	0,30	0,20	0,21
30-20-10	0,43	0,35	0,21	0,22
	ttt eloszláskombináció			
10-15-20	0,53	0,58	0,54	0,53
10-20-30	0,64	0,70	0,69	0,67
20-15-10	0,44	0,47	0,38	0,38
30-20-10	0,50	0,54	0,40	0,41

Az 5. táblában bemutatott és a többi eloszláskombinációval kapcsolatban elvégzett erőelemzések alapján a következő megállapításokat tehetjük különböző elemszámok esetén magas szórásheterogenitási szinten.

i) A rangsorolások egyetlen eloszláskombinációban (ttt) hatékonyabbak, mint a legjobb robusztus paraméteres próba,  $W$ . Minden más eloszláskombinációban  $W$  nagyobb erejű, mint a rangsorolások próbák.  $W$  főlénye különösen magas, ha a magas szórás alacsony csúcossággal párosul (t<sub>nu</sub>, t<sub>uu</sub>, uuu stb.).

ii) Az 5. táblában szereplő négy rangsorolási eljárás közül minden esetben rW2 a legerősebb. A Kulle-féle próbákkal szembeni erőfölénye különösen szembetűnő abban az esetben, amikor az elemszámok és a szórások kapcsolata negatív, vagyis amikor kis minta nagy szórással, nagy minta kis szórással társul.

iii) Szimmetrikus eloszlások esetén az SZTH definíciójában szereplő  $w_i$  súlyok (lásd /10/) irrelevánsak, mert az SZTH fennállása nem függ tőlük. Emiatt KF1, KF2, KG1 és KG2, nullhipotézise közös. Közülük KF1 és KG1 kissé konzervatívabb, mint KF2 és KG2 és ez mérlegelendő előny lehet abban az esetben, amikor ez utóbbiak elsőfajú hibája hajlamos megnőni, azaz (negatív szórás elemszám-kapcsolat esetén). (Lásd a 4. táblát.)

iv) KF2 és KG2 (és hasonlóképpen KF1 és KG1) ereje között bár szisztematikus, de elhanyagolható mértékű különbség figyelhető meg.

Ugyanezen elrendezésekben  $V_e = 0$  esetén, ha az elemszámok azonosak, a bemutatottakhoz hasonló mintázati összefüggések érvényesek (lásd 2. pont a) eset) rW2, KF2 és KG2 végig azonos nagyságrendű hatékonyságával,  $V_e = 0,2$  esetén pedig a  $V_e = 0,4$  szintre vonatkozó összefüggések mutatkoznak, csak kevésbé erőteljes formában.

4. *Folytonos ferde eloszlások.* A két ferde eloszlással végzett szimuláció nagyon hasonló eredményekre vezetett, a következő összefüggésekkel.

a) Az rW2 elsőfajú hibája  $V_e \leq 0,4$  esetén teljesen elfogadható volt: 5 és 10 százalékos szinten egyetlen elsőfokú hibabecslés sem haladta meg szignifikánsan a névleges szint 110 százalékát. Az rW2 érvényessége csak  $V_e = 0,6$  esetén csorbult számottevően. Ez esetben a kapott elsőfajú hibabecslések maximuma 5 százalékos szinten 0,071, 10 százalékos szinten pedig 0,136 volt.

b) A Kulle-féle próbák közül KF1 és KG1, illetve KF2 és KG2 érvényessége gyakorlatilag azonos szintű volt minden elrendezésben. Esetenként KF1 és KF2 elsőfajú hibája egy árnyalattal kedvezőbb, mint KG1-é, illetve KG2-é. Még különböző elemszámok esetén is. Ez azért figyelemre méltó, mert az SZTH-t azonos  $w_i$  súlyokkal állítottam be, ami miatt KG1-nek, illetve KG2-nek kellett volna elméletileg a névleges szintet jobban tartania. Ennek itt feltehetőleg az az oka, hogy a ferde eloszlások az SZTH beállítása után a páronkénti sztochasztikus egyenlőség követelményének is eleget tettek, melynek fennállása esetén nincs különbség az SZTH súlyozott és súlyozatlan fogalma között.

A Kulle-féle próbák elsőfajú hibája érdekes összefüggésben van az elemszámok és a szórások közötti viszonyal. KF2 és KG2 szinte csak akkor mutat a névleges szintet lényegesen meghaladó elsőfajú hibát, ha az  $(n_i, \sigma_i)$  kapcsolat negatív. Például KF2 elsőfajú hibája  $\alpha = 0,05$  esetén az összes ilyen elrendezésben egy kivétellel mindig szignifikánsan nagyobb volt, mint  $1,1\alpha$ , míg más elrendezésekben ilyen megemelkedett elsőfajú hiba csak egyszer fordult elő.

c) Ha a próbák erejét is figyelembe vesszük, akkor a következő kép rajzolódik ki.

i) Azonos elemszámú elrendezésekben alacsony szórásheterogenitás ( $V_e \leq 0,2$ ) esetén legtöbbször KF2, magasabb szórásheterogenitás esetén viszont rW2 ereje tűnik a kedvezőbbnek.

ii) Pozitív  $(n_i, \sigma_i)$  viszony esetén KF2 mutatkozik az SZTH legjobb próbájának, mert rW2 csak akkor erősebb KF2-nél, ha elsőfajú hibája már erőteljesen meghaladja a névleges szintet.

iii) Negatív ( $n_i, \sigma_i$ ) kapcsolat esetén viszont egyértelműen rW2 a legjobb próba, mert kedvező – 0,06-ot sosem meghaladó – elsőfajú hibaszint mellett ereje számottevően nagyobb, mint bármely Kulle-féle próbáé.

6. tábla

A KF2 és az rW2 ereje 5 százalékos szignifikanciaszinten a 10-20-30 elemszámkombináció esetén három szórás heterogenitási szintre vonatkozóan

Próba	Khi-négyzet-eloszlás			Lambda eloszlás		
	$V_e = 0,2$	$V_e = 0,4$	$V_e = 0,6$	$V_e = 0,2$	$V_e = 0,4$	$V_e = 0,6$
KF2	0,61	0,51	0,36	0,27	0,21	0,19
rW2	0,69	0,74	0,68	0,36	0,38	0,39

\*

A jelen tanulmányban független minták összehasonlításának egy új módját mutattuk be. Az „ugyanakkorosság” átlagok segítségével történő definíciójának alternatívájaként bevezettük a sztochasztikus egyenlőség (SZTE) és a sztochasztikus homogenitás (SZTH) fogalmát. Kissé leegyszerűsítve egy  $X$  és egy  $Y$  változóval reprezentált sokaságot akkor mondunk sztochasztikusan egyenlőnek, ha  $X < Y$  ugyanolyan gyakran fordul elő, mint  $X > Y$ . Az SZTH az SZTE általánosítása kettőnél több sokaságra. Az SZTH fennállásának feltétele, hogy az összehasonlított sokaságok egyike se legyen sztochasztikusan kisebb vagy nagyobb, mint a maradék uniója.

Az SZTH tesztelésével kapcsolatban számos paraméteres és nemparaméteres (rangsorolós) próbát vettünk szemügyre, s ezek előnyeit és hátrányait számítógépes szimuláció segítségével elemeztük 10 és 30 közötti nagyságú mintákkal különböző szóródású és eloszlású elrendezésekben. Az elemzések legfontosabb következtetései a következők.

1. Folytonos szimmetrikus eloszlások esetén, amikor az SZTH azonos az elméleti átlagok egyenlőségével, az SZTH nullhipotézisének tesztelésére leginkább a Welch-féle VA (W) a legalkalmasabb. A hagyományos VA ugyan némileg hatékonyabb W-nél abban az esetben, amikor teljesül a normalitás és a szórás homogenitás feltétele, de ezt kis minták esetén aligha lehet megbízhatóan ellenőrizni. A paraméteres W helyett egy esetben célszerű más eljárásról folyamodni: akkor, ha a függő változó eloszlása minden sokaságban sokkal csúcsosabb, mint a normális eloszlás. Ilyenkor a paraméteres próbák (köztük W) kisebb erejűek, mint egyes rangsorolós próbák, amelyek közül egyes esetekben a B. Kulle (1999) által ismertetett eljárás egyik módosítása (KF2), legtöbbször pedig a rang Welch-próba egyik módosítása (rW2) mutatkozik a legmegfelelőbbnek.

2. Ferde eloszlások esetén az SZTH és az átlagok egyenlősége két különböző reláció. Emiatt az SZTH tesztelésére ez esetben csak a rangsorolós eljárások jöhetnek szóba. Közülük ismét csak a KF2 és az rW2 tűnik a legjobbnak. E két eljárás közötti választás erősen függ az elemszámok és a szórások viszonyától. Pozitív ( $n_i, \sigma_i$ ) kapcsolat esetén inkább KF2, negatív ( $n_i, \sigma_i$ ) kapcsolat esetén viszont inkább rW2 az ajánlott eljárás. Azonos nagyságú minták esetén a helyzet felemás: kis szórás heterogenitási szinten inkább KF2, ellenkező esetben inkább rW2 használata a célszerűbb.

Végül megjegyezzük, hogy az SZTH tesztelése igen egyszerűen végrehajtható a MiniStat programcsomag (Vargha–Czigler; 1999) segítségével.

## FÜGGELÉK

## AZ SZTH TESZTELÉSÉRE KULLE ÁLTAL KIDOLGOZOTT STATISZTIKAI PRÓBÁK

Először a mintaelemszámokkal súlyozott ( $w_i = n_i/n$ ), majd a súlyozatlan ( $w_i = 1/h$ ) SZTH tesztelésére alkalmas Kulle-féle eljárást mutatjuk be. Mielőtt azonban rátérnénk a részletekre, bevezetünk néhány jelölést.

Tegyük fel, hogy  $h$  sokaságot akarunk összehasonlítani egy  $X$  változó segítségével úgy, hogy a sokaságokból független mintákat választunk ki. Jelöljük az  $X$  változót a  $j$ -edik sokaságban  $X_j$ -vel, és legyen

$$F_j(x) = (F_j^+(x) + F_j^-(x))/2 \quad (j = 1, \dots, h) \quad /F1/$$

az  $X_j$  változó ún. normalizált eloszlásfüggvénye, ahol  $F_j^+(x) = P(X_j \leq x)$  és  $F_j^-(x) = P(X_j < x)$  az  $X_j$  változó jobbról, illetve balról folytonos eloszlásfüggvénye. Folytonos eloszlások esetén a normalizált eloszlásfüggvény szükségképpen megegyezik a balról és a jobbról folytonos változattal, emiatt a különbségtételnek csak diszkrét eloszlások esetén van jelentősége. Az eloszlásfüggvény ezen alakját már *P. Lévy* (1925) és *W. H. Kruskal* (1952) is megemlíti, majd fontos szerepet kap többek között *E. Brunner*, *M. L. Puri* és *S. Sun* (1995) és *U. Munzel* (1999) munkáiban egyes diszkrét változókkal kapcsolatos rangstatistikák aszimptotikus eloszlásának meghatározásában.

*Az SZTH tesztelése a mintaelemszámokkal arányos súlyozás esetén*

Jelölje az előbbiekkal egyezően  $n = \sum n_j$  az összelemszámot,  $\mathbf{p} = (p_1, \dots, p_h)'$  a  $p_j$  sztochasztikus kezelési hatások  $h$ -dimenziós vektorát,  $\hat{\mathbf{p}} = (\hat{p}_1, \dots, \hat{p}_h)'$  pedig  $\mathbf{p}$ -nek azt a torzítatlan és konzisztens pontbecslését, melynek komponensei a következők:  $\hat{p}_j = (\bar{R}_j - 0,5)/n$ , ahol  $\bar{R}_j$  ( $j = 1, \dots, h$ ) a  $j$ -edik független minta Kruskal–Wallis-próba szerint definiált rangszámátlaga. Kulle (1999. 33. old., 3.24. tétel) bebizonyította, hogy tetszőleges szimmetrikus és idempotens – vagyis projekciós –  $\mathbf{C}$  kontraszt mátrixra (amelyre tehát  $\mathbf{C}' = \mathbf{C}$  és  $\mathbf{C}\mathbf{C} = \mathbf{C}$ ) a

$$H_0: \mathbf{C}\mathbf{p} = 0 \quad /F2/$$

nullhipotézis fennállása esetén igaz a következő két állítás.

a) A  $\sqrt{n}(\hat{\mathbf{p}} - \mathbf{p})$  valószínűségi vektorváltozó aszimptotikusan azonos eloszlású egy olyan aszimptotikusan normális  $\mathbf{Z}$  vektorváltozóval, amelynek komponensei a következők:

$$Z_j = \sqrt{n} \left( \int H d\hat{F}_j - \int F_j d\hat{H} \right) \quad (1 \leq j \leq h). \quad /F3/$$

E formulában  $\hat{F}_j$  az /F1/ formulával definiált  $F_j$ ,  $\hat{H}$  pedig ezen  $F_j$ -k  $H = \sum w_j F_j$  súlyozott átlagával definiált empirikus eloszlásfüggvénye, ahol  $w_j = n_j/n$ . (Matematikailag igazolható, hogy e jelölések figyelembevételével a sztochasztikus kezelési hatások felírhatók a  $p_j = \int H dF_j$  alakban is.)

b)  $\mathbf{V}$ -vel jelölve a  $\mathbf{Z}$  vektorváltozó kovarianciamátrixát, ha vesszük  $\mathbf{V}$ -nek az /F14/, /F15/ formulákkal meghatározott  $\hat{\mathbf{V}}$  becslését, akkor az

$$F_K = \frac{n\hat{\mathbf{p}}'\mathbf{C}\hat{\mathbf{p}}}{tr(\hat{\mathbf{C}\mathbf{V}})} \quad /F4/$$

próbatasztika aszimptotikusan  $\chi^2(f)/f$  eloszlást követ a következő módon becsülhető  $f$  szabadságfokkal:

$$\hat{f} = \frac{(tr(\hat{\mathbf{C}\mathbf{V}}))^2}{tr(\hat{\mathbf{C}\mathbf{V}\mathbf{C}\mathbf{V}})} \quad /F5/$$

(e formulákban  $tr$  a vonatkozó mátrix nyomát (trace) jelöli).

Az előbbi általános formában megfogalmazott tétel alapján az SZTH nullhipotézise a következő  $c_{ij}$  együtthatókkal definiált  $\mathbf{C}$  kontraszt mátrix segítségével tesztelhető:

$$c_{ij} = \frac{h-1}{h} \quad (1 \leq j \leq h), \quad c_{ij} = -\frac{1}{h} \quad (1 \leq i \neq j \leq h). \quad /F6/$$

Egyszerűen belátható, hogy  $\mathbf{C}$  projekciós mátrix, és hogy  $\mathbf{C}\mathbf{p} = 0$  maga után vonja az SZTH fennállását. A próba végrehajtásához szükséges  $F_K$  és  $\hat{f}$  mennyiséget a következőképpen lehet kiszámítani.

1.  $F_K$  számlálója az alábbi formula segítségével adódik:

$$n\hat{\mathbf{p}}'\mathbf{C}\hat{\mathbf{p}} = \frac{1}{n} \sum_{j=1}^h (\bar{R}_j - \bar{R})^2, \quad /F7/$$

ahol  $\bar{R}_j$  a  $j$ -edik független minta rangszámátalaga ( $1 \leq j \leq h$ ),  $\bar{R}$  pedig ezen  $\bar{R}_j$  mennyiségek súlyozatlan átlaga.

2.  $F_K$  és az /F5/ formulával felírt  $\hat{f}$  szabadságfok kiszámításához szükség van a  $\mathbf{C}\hat{\mathbf{V}}$  mátrix nyomára (lásd /F4/), melyre igaz az alábbi összefüggés:

$$\text{tr}(\mathbf{C}\hat{\mathbf{V}}) = \sum_{i=1}^h \sum_{j=1}^h c_{ij} \hat{v}_{ji} = \frac{h-1}{h} \sum_{i=1}^h \hat{s}_i^2 - \frac{1}{h} \sum_{i=1}^h \sum_{j \neq i}^h \hat{v}_{ij}. \quad /F8/$$

Ehhez azonban elsőként a szimmetrikus  $\hat{\mathbf{V}}$  empirikus kovarianciamátrix főátlójának  $\hat{s}_i^2$  és főátlón kívüli  $\hat{v}_{ji} = \hat{v}_{ij}$  elemeit kell meghatározni, melyhez a következő lépések szükségesek.

3. Definiáljunk először is mindegyik  $X_{ik}$  ( $i = 1, \dots, h; k = 1, \dots, n_i$ ) megfigyeléshez egy ún.  $W_{ik}(j)$  részhelyértéket minden  $j \neq i, 1 \leq j \leq h$  indexre a következő formula segítségével:

$$W_{ik}(j) = \sum_{r=1}^{n_j} c(X_{ik} - X_{jr}), \quad /F9/$$

amelyben a  $c(x)$  számlálófüggvény értéke minden  $x$  valós számra 1, 0,5, illetve 0, attól függően, hogy  $x > 0$ ,  $x = 0$  vagy pedig  $x < 0$ . Köznapi szavakkal  $W_{ik}(j)$  azon esetek száma, ahányszor az  $X_{ik}$  megfigyelés nagyobb, mint egy  $j$ -edik mintabeli megfigyelés, továbbá azon esetek számának fele, amikor  $X_{ik}$  egyenlő egy  $j$ -edik mintabeli megfigyeléssel. Minden  $X_{ik}$  megfigyelésre a  $W_{ik}(j)$  részhelyértékek  $W_{ik}$  összegét az  $X_{ik}$  megfigyeléshez tartozó helyértéknek nevezzük:

$$W_{ik} = \sum_{j \neq i}^h W_{ik}(j). \quad /F10/$$

A helyérték itt bemutatott fogalma *M. A. Fligner* és *G. E. Policello* (1981) két független minta esetében bevezetett hasonló fogalmának közvetlen általánosítása kettőnél több független mintára. Megjegyezzük, hogy a Kruskal–Wallis-próba szerinti rangsorolás esetén kapott  $R_{ik}$  rangérték minden  $(i, k)$  indexpárra a  $W_{ik}$  helyérték és az  $X_{ik}$  megfigyelés  $i$ -edik mintán belüli rangsorolásával kapott rangérték sima összege.

4. A  $W_{ik}$  helyértékek és a  $W_{ik}(j)$  részhelyértékek segítségével definiáljuk most az alábbi mennyiségeket:

a) a  $W_{ik}$  ( $k = 1, \dots, n_i$ ) helyértékek  $i$ -edik mintán belüli tapasztalati varianciája osztva  $n^2$ -tel:

$$\hat{\sigma}_i^2 = \frac{1}{n^2} \text{Var}(W_{ik}) \quad (1 \leq i \leq h); \quad /F11/$$

b) a  $W_{ik}(j)$  ( $k = 1, \dots, n_i$ ) részhelyértékek  $i$ -edik mintán belüli tapasztalati varianciája tetszőleges rögzített  $i$ ,  $j$  ( $i \neq j$ ) indexekre osztva  $n^2$ -tel:

$$\hat{\tau}_{i,j}^2 = \frac{1}{n^2} \text{Var}(W_{i,j}) \quad (1 \leq i \neq j \leq h); \quad /F12/$$

c) A  $W_{ik}(r)$ ,  $W_{ik}(s)$  ( $k = 1, \dots, n_i$ ) részhelyértékek tapasztalati kovarianciája tetszőleges rögzített  $i$ ,  $r$ ,  $s$  ( $r \neq i$ ,  $s \neq i$ ) indexekre osztva  $n^2$ -tel:

$$\hat{\alpha}_i(r,s) = \frac{1}{n^2} \text{Cov}(W_{i,r}, W_{i,s}) \quad (1 \leq r \neq i, s \neq i \leq h). \quad /F13/$$

5. Ezután Kulle (1999. 17. és 20. old.) alapján a  $\hat{V}$  tapasztalati kovarianciamátrix  $\hat{s}_i^2$  és  $\hat{v}_{ij}$  együtthatóit a következőképpen kaphatjuk meg az előbbieken definiált mennyiségek segítségével:

$$\hat{s}_i^2 = \hat{v}_{ii} = \frac{n}{n_i} \hat{\alpha}_i^2 + \frac{n}{n_i^2} \sum_{r \neq i}^h n_r \hat{\tau}_{r,i}^2 \quad /F14/$$

és  $i \neq j$  esetén

$$\hat{v}_{ij} = \frac{n}{n_i n_j} \left( -n_i \hat{\tau}_{i,j}^2 - n_j \hat{\tau}_{j,i}^2 - \sum_{k \neq i,j}^h (n_i \hat{\alpha}_i(j,k) + n_j \hat{\alpha}_j(i,k) - n_k \hat{\alpha}_k(i,j)) \right). \quad /F15/$$

6. Végül az  $\hat{f}$  szabadságfokot definiáló /F5/ tört nevezőjét a következőképpen nyerjük:

$$\text{tr}(\mathbf{C}\hat{V}\mathbf{C}\hat{V}) = \sum_{i=1}^h \sum_{j=1}^h m_{ij} m_{ji}, \quad /F16/$$

melyben a  $\mathbf{C}\hat{V}$  mátrix  $m_{ij}$  ( $1 \leq i, j \leq h$ ) elemeire a következő összefüggés érvényes:

$$m_{ij} = \frac{h-1}{h} \hat{v}_{ij} - \frac{1}{h} \sum_{k \neq i}^h \hat{v}_{kj}. \quad /F17/$$

Az SZTH tesztelése azonos súlyozás esetén

Azonos súlyok esetén az SZTH tesztelése Kullének egy másik tételén alapul (Kulle; 1999. 62. old., 4.21. tétel), mely a következőképpen szól. Legyen  $\mathbf{p} = (p_1, \dots, p_h)'$  a  $p_j$  sztochasztikus kezelési hatások  $h$ -dimenziós vektora (ez esetben azonban a /10/ formulában azonos  $w_j = 1/h$  súlyokat alkalmazva),  $\hat{\mathbf{p}} = (\hat{p}_1, \dots, \hat{p}_h)'$  pedig  $\mathbf{p}$ -nek az /F29/ formulával meghatározott pontbecslése. Kulle bebizonyította, hogy tetszőleges projekciós  $\mathbf{C}$  kontraszt mátrixra a  $H_0: \mathbf{C}\mathbf{p} = \mathbf{0}$  nullhipotézis fennállása esetén igaz a következő két állítás:

a) a  $\sqrt{n}(\hat{\mathbf{p}} - \mathbf{p})$  valószínűségi vektorváltozó aszimptotikusan azonos eloszlású egy olyan aszimptotikusan normális  $\mathbf{Y}$  vektorváltozóval, amelynek komponensei a következők:

$$Y_j = \sqrt{n} \left( G d \hat{F}_j - \int F_j d \hat{G} \right) \quad (1 \leq j \leq h), \quad /F18/$$

e formulában  $\hat{F}_j$  az /F1/ formulával definiált  $F_j$ ,  $\hat{G}$  pedig ezen  $F_j$ -k  $G = h^{-1} \sum_j F_j$  súlyozatlan átlagával definiált eloszlásfüggvény empirikus eloszlásfüggvénye (matematikailag igazolható, hogy e jelölések figyelembevételével a sztochasztikus kezelési hatások ez esetben a  $p_j = \int G d F_j$  alakban is felírhatók);

b)  $\mathbf{S}$ -sel jelölve az  $\mathbf{Y}$  vektorváltozó kovarianciamátrixát, ha vesszük  $\mathbf{S}$ -nek az /F6/ és /F7/ formulával meghatározott  $\hat{\mathbf{S}}$  becslését, akkor a

$$G_K = \frac{n\hat{\mathbf{p}}'\hat{\mathbf{C}}\hat{\mathbf{p}}}{\text{tr}(\hat{\mathbf{C}}\hat{\mathbf{S}})} \quad /F19/$$

próbatesztstatistika aszimptotikusan  $\chi^2(g)/g$  eloszlást követ a következőképpen becsülhető  $g$  szabadságfokkal:

$$\hat{g} = \frac{(\text{tr}(\hat{\mathbf{C}}\hat{\mathbf{S}}))^2}{\text{tr}(\hat{\mathbf{C}}\hat{\mathbf{S}}\hat{\mathbf{C}}\hat{\mathbf{S}})}, \quad /F20/$$

és ez esetben az SZTH nullhipotézise szintén az /F6/-tal megadott  $\mathbf{C}$  kontraszt mátrix segítségével tesztelhető.

A próba végrehajtásához szükséges  $G_K$  és  $\hat{g}$  mennyiség kiszámítása a következő lépésekben történhet.

1. Ebben az esetben definiáljuk először a részhelyértékeket a következőképpen. Az  $i$ -edik minta tetszőleges  $X_{ik}$  megfigyelése esetén tetszőleges rögzített  $j \neq i$ ,  $1 \leq j \leq h$  indexre legyen a

$$\tilde{W}_{ik}(j) = \frac{\tilde{n}}{n_j} \sum_{r=1}^{n_j} c(X_{ik} - X_{jr}) = \frac{\tilde{n}}{n_j} W_{ik}(j), \quad /F21/$$

mennyiség az  $X_{ik}$ -hoz tartozó  $j$ -edik részhelyérték, ahol  $\tilde{n} = [h/(n_i)^{-1}]^{-1}$  az  $n_i$  elemszámok harmonikus átlaga. A mintaelemszámokkal súlyozott esettel analóg módon definiáljuk az  $X_{ik}$ -hoz tartozó helyértéket is:

$$\tilde{W}_{ik} = \sum_{j \neq i} W_{ik}(j). \quad /F22/$$

2. A helyértékek és részhelyértékek segítségével képezzük az alábbi mennyiségeket:

$$\hat{\sigma}_i^2 = \frac{1}{(h\tilde{n})^2} \text{Var}(\tilde{W}_i) \quad (1 \leq i \leq h) \quad /F23/$$

$$\hat{\tau}_{i,j}^2 = \frac{1}{(h\tilde{n})^2} \text{Var}(\tilde{W}_i(j)) \quad (1 \leq i \neq j \leq h) \quad /F24/$$

$$\hat{\alpha}_i(r,s) = \frac{1}{(h\tilde{n})^2} \text{Cov}(\tilde{W}_i(r), \tilde{W}_i(s)) \quad (1 \leq r \neq i, s \neq i \leq h). \quad /F25/$$

3. Ezután Kulle (1999. 49. és 53. old.) alapján az  $\hat{\mathbf{S}}$  tapasztalati kovarianciamátrix együtthatóit a következőképpen kaphatjuk meg az előbb definiált mennyiségek segítségével:

$$\hat{s}_{ii}^2 = \hat{s}_{ii} = \frac{N}{n_i} \hat{\sigma}_i^2 + \frac{N}{n_i^2} \sum_{r \neq i} n_r \hat{\tau}_{ri}^2, \quad /F26/$$

és  $i \neq j$  esetén

$$\hat{s}_{ij} = -\frac{n}{n_i} \hat{\tau}_{i,j}^2 - \frac{n}{n_j} \hat{\tau}_{j,i}^2 - n \cdot \sum_{k \neq i,j} \left( \frac{\hat{\alpha}_i(j,k)}{n_i} + \frac{\hat{\alpha}_j(i,k)}{n_j} - \frac{\hat{\alpha}_k(i,j)}{n_k} \right). \quad /F27/$$

4. Az /F19/ formulával definiált  $G_K$  próbataszitika számlálóját a következőképpen kapjuk:

$$n\hat{\mathbf{p}}' \mathbf{C} \hat{\mathbf{p}} = n \sum_{j=1}^h \hat{p}_j^2 - \frac{n}{h} \left( \sum_{j=1}^h \hat{p}_j \right)^2, \quad /F28/$$

ahol a  $p_j$  sztochasztikus kezelési hatások  $\hat{p}_j$  becslései a következő összefüggés segítségével adódnak:

$$\hat{p}_j = 0,5 + \sum_{r \neq j} \sum_{k=1}^h \frac{\tilde{W}_{jk}(r)}{hn_j n_r}. \quad /F29/$$

5. A  $G_K$  próbataszitika nevezőjét a következő formula segítségével határozhatjuk meg:

$$tr(\hat{\mathbf{C}}\hat{\mathbf{S}}) = \sum_{i=1}^h \sum_{j=1}^h c_{ij} \hat{s}_{ji} = \frac{h-1}{h} \sum_{i=1}^h \hat{s}_i^2 - \frac{1}{h} \sum_{i=1}^h \sum_{j \neq i} \hat{s}_{ij}. \quad /F30/$$

6. Végül a  $\hat{g}$  szabadságfokot az /F16/ és az /F17/ formula segítségével nyerjük oly módon, hogy bennük a  $\hat{v}_{ij}$  együtthatókat a megfelelő  $\hat{s}_{ij}$  együtthatókkal helyettesítjük.

#### IRODALOM

- AKRITAS, M. G. – ARNOLD, S. F. (1994): Nonparametric hypotheses for factorial designs: Multivariate repeated measures designs. *Journal of the American Statistical Association*, 89. évf. 336–343. old.
- AKRITAS, M. G. – ARNOLD, S. F. – BRUNNER, E. (1997): Nonparametric hypotheses and rank statistics for unbalanced factorial designs. *Journal of the American Statistical Association*, 92. évf. 258–265. old.
- BRADLEY, J. V. (1978): Robustness? *British Journal of Mathematical and Statistical Psychology*, 34. évf. 144–152. old.
- BROWN, M. B. – FORSYTHE, A. B. (1974): The small sample behavior of some statistics which test the equality of several means. *Technometrics*, 16. évf. 129–132. old.
- BRUNNER, E. – MUNZEL, U. (2000): The nonparametric Behrens-Fisher problem: Asymptotic theory and a small-sample approximation. *Biometrical Journal*, 42. évf. 17–25. old.
- BRUNNER, E. – PURI, M. L. (2001): Nonparametric methods in factorial designs. *Statistical Papers*, 42. évf. 1–52. old.
- BRUNNER, E. – PURI, M. L. – SUN, S. (1995): Nonparametric methods for two-sample designs with application to multiclinic trials. *Journal of the American Statistical Association*, 90. évf. 1004–1014. old.
- CLIFF, N. (1996): *Ordinal methods for behavioral data analysis*. Lawrence Erlbaum Associates, Mahwah, New Jersey.
- CLINCH, J. J. – KESELMAN, H. J. (1982): Parametric alternatives to the analysis of variance. *Journal of Educational Statistics*, 7. évf. 207–215. old.
- COHEN, J. (szerk.) (1977): *Statistical power analysis for the behavioral sciences*. Academic Press, New York.
- CONOVER, W. J. – IMAN, R. L. (1981): Rank transformations as a bridge between parametric and nonparametric statistics. *The American Statistician*, 35. évf. 124–129. old.
- DIXON, W. J. (1990): *BMDP Statistical software manual*. University of California Press, Berkeley.
- FLIGNER, M. A. – POLICELLO II, G. E. (1981): Robust rank procedures for the Behrens-Fisher problem. *Journal of the American Statistical Association*, 76. évf. 323–327. old.
- JAMES, G. S. (1951): The comparison of several groups of observations when the ratios of the population variances are unknown. *Biometrika*, 38. évf. 324–329. old.
- KRUSKAL, W. H. (1952): A nonparametric test for the several sample problem. *The Annals of Mathematical Statistics*, 23. évf. 525–540. old.
- KRUSKAL, W. H. – WALLIS, W. A. (1952): Use of ranks. *Journal of the American Statistical Association*, 47. évf. 583–621. old.
- KULLE, B. (1999): *Nichtparametrisches Behrens-Fisher-Problem im Mehr-Stichprobenfall*. Doctoral Thesis. Institut für Mathematische Stochastik der Georg-August-Universität, Göttingen.
- LÉVY, P. (1925): *Calcul des probabilités*. Gauthiers-Villard, Paris.
- MANN, H. B. – WHITNEY, D. R. (1947): On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, 18. évf. 50–60. old.
- MICCERI, T. (1989): The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin*, 105. évf. 156–166. old.
- MUNZEL, U. (1999): Linear rank score statistics when ties are present. *Statistical Probability Letters*, 41. évf. 389–395. old.
- NOETHER, G. E. (1967): *Elements of nonparametric statistics*. Wiley, New York.
- ONGHENA, P. (1993): A theoretical and empirical comparison of mainframe, microcomputer, and pocket calculator pseudorandom number generators. *Behavior Research Methods, Instruments, & Computers*, 25. évf. 384–395. old.



- PRESS, W. H. – FLANNERY, B. P. – TEUKOLSKY, S. A. – VETTERLING, W. T. (1989): *Numerical recipes in Pascal. The art of scientific computing*. Cambridge University Press, Cambridge.
- RAMBERG, J. S. – TADIKAMALLA, P. R. – DUDEWICZ, E. J. – MYKYTKA, E. F. (1979): A probability distribution and its uses in fitting data. *Technometrics*, 21. évf. 201–209. old.
- SCHEFFÉ, H. (1959): *The analysis of variance*. Wiley, New York.
- VARGHA A. (1981): *Pszichológiai statisztika gyakorlat II. Egységes jegyzet*. Tankönyvkiadó, Budapest.
- VARGHA A. (1996): Az egymintás  $t$ -próba érvényessége és javíthatósága. *Magyar Pszichológiai Szemle*, 52. évf. 317–345. old.
- VARGHA A. (1999): Két csoport összehasonlítása nemparaméteres statisztikai eljárások segítségével. *Magyar Pszichológiai Szemle*, 54. évf. 567–589. old.
- VARGHA A. (2000a): *Matematikai statisztika pszichológiai, nyelvészeti és biológiai alkalmazásokkal*. Pólya Kiadó, Budapest.
- VARGHA A. (2000b): Két pszichológiai populáció sztochasztikus egyenlőségének ellenőrzésére alkalmas statisztikai próbák összehasonlító vizsgálata. *Magyar Pszichológiai Szemle*, 55. évf. 253–281. old.
- VARGHA A. – CZIGLER B. (1999): *A MiniStat statisztikai programcsomag, 3.2 verzió*. Pólya Kiadó, Budapest.
- VARGHA, A. – DELANEY, H. D. (1998): The Kruskal-Wallis test and stochastic homogeneity. *Journal of Educational and Behavioral Statistics*, 23. évf. 170–192. old.
- VARGHA, A. – DELANEY, H. D. (2000): A critique and improvement of the CL common language effect size statistics of McGraw and Wong. *Journal of Educational and Behavioral Statistics*, 25. évf. 101–132. old.
- VINCZE I. (1968): *Matematikai statisztika ipari alkalmazásokkal*. Műszaki Könyvkiadó, Budapest.
- WELCH, B. L. (1951): On the comparison of several mean values: An alternative approach. *Biometrika*, 38. évf. 330–336. old.
- WILCOX, R. R. (1996): *Statistics for the social sciences*. Academic Press, San Diego, New York.
- ZIMMERMAN, D. W. – ZUMBO, B. D. (1992): Parametric alternatives to the Student  $t$  test under violation of normality and homogeneity of variance. *Perceptual and Motor Skills*, 74. évf. 835–844. old.
- ZIMMERMAN, D. W. – ZUMBO, B. D. (1993): Rank transformations and power of the Student  $t$  test and Welch  $t'$  test for non-normal populations with unequal variances. *Canadian Journal of Experimental Psychology*, 47. évf. 523–539. old.

## SUMMARY

Non-quantitative ordinally scaled variables occur with a high frequency in social and behavioral sciences. In the case of these types of variables the comparison of the value levels by means of the expected values may be questionable or misleading.

The present paper asserts that stochastic comparison, with the key terms of stochastic equality (STE) and stochastic homogeneity (STH), is a powerful alternative to conventional parametric comparison methods. Two populations are said to be stochastically equal with respect to a variable  $X$  if  $P(X_1 > X_2) = P(X_1 < X_2)$ , where  $X_1$  and  $X_2$  are randomly chosen independent observations from the two populations respectively. STH is a direct generalization of STE by means of which one can compare more than two populations.

After an explanation of the related theoretical terms the paper presents a simulation study where the advantages and disadvantages of 15 statistical tests of STH are scrutinized.

# TÉRÖKONOMETRIA\*

VARGA ATTILA

A térökonometria (spatial econometrics) az ökonometria azon részterülete, mely a keresztmetszeti és paneladatokra épülő regressziós modellekben a térbeli interakció (térbeli autokorreláció) és a térbeli struktúra (térbeli heterogenitás) által generált ökonometriai problémák kezelésével foglalkozik. A tanulmány elsősorban a térökonometria szemlélet- és módszertanbeli sajátosságainak magyar nyelven való bemutatására vállalkozik s korántsem törekszik a tudományág teljes körű elemzésére. Ennek megfelelően a térökonometriai modellek közül a széleskörűen alkalmazott lineáris modellek állnak a középpontban. A tanulmány a térökonometriai modellezésének, a regressziós modellek becslési eljárásainak, valamint a térbeli autokorreláció tesztelésének problémáit mutatja be.

TÁRGYSZÓ: Ökonometria; Területi elemzések; Autokorreláció.

A térökonometria az ökonometria azon részterülete, mely a keresztmetszeti és paneladatokra épülő regressziós modellekben a térbeli interakció (térbeli autokorreláció) és a térbeli struktúra (térbeli heterogenitás) által generált ökonometriai problémák kezelésével foglalkozik (*Paelinck–Klaassen*; 1979, *Anselin*; 1988a). A térökonometria bizonyos hasonlóságot mutat a térstatisztikával, melyet széleskörűen alkalmaznak bizonyos fizikai tudományokban (mint például a statisztikai mechanika, az ökológia vagy az epidemiológia), ám ugyanakkor különbözik is tőle s ez a különbség a statisztikai és ökonometriai megközelítés alapvető eltéréséből következik. Ugyanis míg az ökonometriai elemzés jellege, illetve az alkalmazott specifikációk mindig közgazdaság-elméleti modellek által meghatározottak, a statisztika kiindulópontja elsősorban az adatokban van, s így közelítőmódját korántsem jellemzi egy ekképpen értelmezett szigorúan a priori elméleti meghatározottság.

*Paelinck és Klaassen* belga ökonometrikusok a térökonometria alapjait eredetileg a keresztmetszeti adatokat használó regionális és multiregionális modellekben fellépő térbeli autokorreláció kezelésére dolgozták ki az 1970-es évek végén (*Paelinck–Klaassen*; 1979). A térökonometria modelljeinek továbbfejlesztése, illetve azok alkalmazása kezdetben elsősorban a közgazdaságtan specializált, kifejezetten térbeli problémákkal foglalkozó terü-

\* A szerző ezúton is szeretné köszönetét kifejezni *Luc Anselinnek*, a térökonometria vezető nemzetközi kutatójának, akit professzoraként, PhD disszertációs témavezetőjeként, majd kollégaként és több tanulmányban szerzőtársaként tisztelhet. A tanulmány korábbi változatairól készített alapos véleményekért *Hunyadi Lászlót*, *Nemes Nagy Józsefet*, *Pintér Józsefet* és *Rappai Gábor*t illeti köszönet. Az esetleges hibákért természetesen a szerző viseli a felelősséget.

letei – úgy, mint a regionális közgazdaságtudomány, a városgazdaságtan, a gazdaságföldrajz vagy az ingatlan-gazdaságtan – határain belül zajlott. Az 1990-es évek új fejleménye, hogy a térökonometria módszereinek felhasználására egyre több példával találkozhatunk a közgazdaságtan olyan ágaiban is, ahol a folyamatok térbeli vonatkozásainak kutatása hagyományosan nem tartozott az érdeklődés homlokterébe, így a világgazdaságtanban, az államháztartástanban, a munka gazdaságtanában, az agrárgazdaságtanban, vagy a technológiai fejlődés gazdaságtanában (lásd például *Anselin–Varga–Acs*; 1997, *Aten*; 1996, *Bell–Bockstael*; 1999, *Case*; 1991, *Case–Rosen–Hines*; 1993, *Holtz–Eakin*; 1994, *Murdoch–Rahmatian–Thayer*; 1993, *Murdoch–Sandler–Sargent*; 1997, *Nelson–Hellerstein*; 1997, *Topa*; 1996, *Varga*; 1998). Mindezek mellett az utóbbi években az irányadó ökonometriai irodalomban is megjelent több olyan tanulmány, mely egyes modellspecifikációk, teszt-statisztikák és becslési eljárások térben kódolt adatokra való kidolgozását célozza (lásd például *Conley*; 1996, *Driscoll–Kray*; 1998, *Pinkse–Slade*; 1998).

A térökonometriai alkalmazások iránt az 1990-es évektől megéleződött közgazdaságtani érdeklődés két tényező hatására vezethető vissza (*Anselin*; 2001). Az egyik tényezőként az a gazdaságelméletben jelentkező új keletű irányzat említhető, mely (szakítva a közgazdaságtanban mindmáig uralkodó atomisztikus, elszigetelt módon döntést hozó gazdasági szereplő képével) a hangsúlyt az egyének között kiépülő (innovációs, termelési, értékesítési) hálózatok, a gazdasági szereplők térbeli koncentrációjából fakadó extern hatások vagy a tudás-spilloverek formájában működő interakciók tanulmányozására helyezi át. Ez az új szemlélet jelentkezik például az Aoki-féle „új makroökonómiában” (*Aoki*; 1994, 1996), a Romer-Lucas-féle „új (endogén) növekedélméletben” (*Romer*; 1990, *Lucas*; 1988), a Krugman-féle „új gazdaságföldrajzban” (*Krugman*; 1991, *Fujita–Krugman–Venables*; 1999), vagy az „innovációs rendszerek” Nelson-Lundvall-féle irányzatában (*Nelson*; 1993, *Lundvall*; 1992).<sup>1</sup> A térökonometria iránti megnövekedett érdeklődés tehát egyrészt az empirikus térbeli modellek kezelésére képes ökonometriai módszerek iránti igényből fakad. Az alkalmazások szaporodását másrészt a térinformatika (Geographical Information System – GIS) gyors terjedésével párhuzamosan elsősorban az Egyesült Államokban megjelent nagyméretű, térben kódolt elektronikus adatbázisok (vagyis olyan adatbázisok, melyek a megfigyelési egységek – megyék, városok, kerületek stb. – térbeli elhelyezkedésére vonatkozóan is tartalmaznak információt) ökonometriai elemzésének igénye magyarázza. A térbeli adatoknak ugyanis meghatározó sajátossága a térbeli autokorreláció, melynek kezelésére az idősoros vagy a paneladatok elemzésénél általában használt ökonometriai technikák nem alkalmasak.

Tanulmányomban elsősorban a térökonometria szemlélet- és módszertanbeli sajátosságainak magyar nyelven való bemutatására vállalkozom s korántsem törekszem a tudományág teljes körű elemzésére. Így a térökonometriai modellek közül a széleskörűen alkalmazott lineáris modellek állnak a középpontban, míg a kutatás aktív területeiként számon tartott egyéb ökonometriai problémák (például a térbeli panelmodellek, vagy a kategorizált adatok térökonometriai elemzése<sup>2</sup>) részletes bemutatása kívül marad a ta-

<sup>1</sup> Az új növekedésméletről, az innovációs rendszerek irányzatáról és az új gazdaságföldrajzról részletesebben magyarul lásd *Acs és Varga* (2001) tanulmányát.

<sup>2</sup> A térbeli panelmodellel kapcsolatos módszertani kérdésekről lásd például *Anselin* (1988a), *Baltagi–Li* (1999), a térbeli probitmodellel kapcsolatban pedig *Case* (1992), *McMillen* (1992), *Bolduc–Fortin–Gordon* (1997), *Heagerty–Lele* (1998), *Pinkse–Slade* (1998), *Beron–Vijverberg* (1999).

nulmány keretein.<sup>3</sup> A következő rész a tér ökonometriai modellezése problémájának felvázolását és az erre adott térökonometriai megoldást mutatja be. A harmadik rész a térbeli autokorreláció statisztikai teszteléséről, míg a negyedik a lineáris modellekről és azok becslési eljárásairól ad áttekintést. A térbeli autokorrelációval összefüggő specifikációs tesztek ismertetése az ötödik rész tárgya. A tanulmány néhány záró gondolattal fejeződik be.

### A TÉR ÖKONOMETRIAI REPREZENTÁCIÓJA

A térbeli adatok adekvát ökonometriai elemzésének alapfeltétele a megfigyelési egységek közötti térbeli relációk megfelelő módon való leképezése. A tér (közgazdasági szempontból) kétdimenziós jellegéből következően ugyanis azok az ökonometriai megoldások, melyek az „egydimenziós” idősoros adatokra épülő elemzéseknel megfelelőképpen használhatók, a térbeli problémák megragadására nem feltétlenül alkalmasak. A térbeli keresztmetszeti adatbázisok összeállításánál gyakran használt megoldás például az, amikor a megfigyelési egységek betűrendben vagy statisztikai kódszámok alapján kerülnek sorbarendezésre. Mint a későbbiekben látni fogjuk, bizonyos esetekben az idősoros elemzések ezen „egydimenziós” technikájának térbeli adaptációja elfogadható, az esetek többségében viszont nem eredményez megfelelő megoldást.

Idősoros elemzéskor a megfigyelési egységeket időbeli előfordulásuk alapján rendezzük, így az egységek sorrendje megfelel azok időbeli sorrendjének. Az időbeli folyamatok egyirányúsága miatt a valóságnak ez a leképezése teljesen megfelelő. Ez a leképezés lehetővé teszi, hogy mind a megfigyelési egységek egymás közötti függőségéből (időbeli korreláció), mind azok valószínűség-eloszlás függvényeinek különbözőségeiből (időbeli heterogenitás) eredő ökonometriai problémákat kezelni tudjunk. Az időbeli megfigyelések tehát statisztikai–ökonometriai szempontból a legjobban az „idővonal” egymást követő pontjai-ként képezhetőek le. Térbeli adatok esetén a térkép a valóság hasonlóan absztrakt reprezentációja. Azok a módszerek, melyek alkalmasak az egydimenzióban mozgó időbeli folyamatok elemzésére, nem feltétlenül alkalmazhatók egy kétdimenziós jelenség vizsgálatára. Ha a térbeli megfigyelési pontok egymástól függetlenek, akkor e pontok egymáshoz viszonyított helyzetének ismerete nem indokolja speciális ökonometriai módszerek alkalmazását. Akkor viszont, ha a megfigyelési egységeket egymás közötti korreláció jellemzi, vagy azok egymástól igen eltérő jellegzetességekkel bírnak (térbeli heterogenitás), a relatív térbeli pozíciók figyelmen kívül hagyása téves eredményekre vezethet.

Mind a térstatisztikusok mind a geográfusok evidenciaként kezelik azt, hogy a térbeli adatok egyik meghatározó jellemzője azok heterogenitásra és korrelációra való hajlama. A térbeli heterogenitás a térbeli egységek sajátosságainak, másoktól való különbözőségeinek következménye. Városok, megyék stb. ugyanis igen jelentősen különbözhetnek egymástól bizonyos gazdaságilag meghatározó jellemzők szempontjából. Így például jelentős különbségek tapasztalhatóak a megfigyelési egységek között az ott élő munkaerő minőségében, számában, az ott található természeti erőforrások mennyiségében, vagy akár a kutatás–fejlesztési kiadásokban.

A megfigyelési egységek térbeli függősége (vagy térbeli autokorrelációja – „spatial autocorrelation”) inkább szabálynak, mint kivételnek tekinthető. A térstatisztika vezető

<sup>3</sup> A jelenleginél jóval részletesebb elemzésekhez a következő áttekintő munkákat ajánlom az olvasó figyelmébe: *Anselin* (1988a, 2001), *Anselin–Florax* (1995), *Anselin–Rey* (1997), *LeSage* (1999).

nemzetközi tekintélye, *Cressie* véleménye szerint térbeli adatok esetén „a dependencia minden irányban jelen van és minél szétszórtabban helyezkednek el a megfigyelési egységek, annál gyengébb a függőség” (*Cressie*; 1993, 3. old.). A térbeli dependencia lényegében függvényszerű relációt jelent ugyanazon változó különböző helyeken mért értékei között. Míg az adatok időbeli függősége esetén a két megfigyelési egység közötti reláció egyirányú, addig a térbeli függőség kétirányú reláció kettőnél több térbeli egység között, vagyis egy változó bármely lokációban mért értéke nemcsak, hogy hat a többi lokációkban mért értékekre, de ugyanakkor maga is determinált az egyéb lokációkban mért értékek által.

Hasonló értékek csoportosulása a térben (az ún. klaszterek kialakulása) a megfigyelési egységekben mért értékek között a pozitív térbeli autokorreláció létrejöttét eredményezheti. Példaként hozható a nagyvárosok tehetősebbek vagy éppen szegények lakta negyedeinek csoportulása a térben vagy a világ nagy innovációs klasztereinek jelensége, ahol az egyes klaszterek számos, a technológiai innovációban sikeres nagyváros térbeli „sűrűsödésének” eredményei. Negatív térbeli autokorreláció esetén az egyes megfigyelési egységek szomszédságában azoktól különböző értékeket mérünk, ami az értékek egyfajta saktáblaszerű térbeli eloszlását eredményezi. A kétfajta autokorreláció közül a pozitív autokorreláció jelenségét jóval könnyebb megérteni, míg a negatív autokorrelációra nem minden esetben tudunk intuitív magyarázatot találni.

Az adatokban észlelt pozitív térbeli autokorreláció két eltérő okkal magyarázható: az egyik technikai jellegű, míg a másik szubsztantív, vagyis a társadalmi interakciók természetéből következik. A technikai ok a számbavétel esetleges hibáiból adódik. Térbeli adatok aggregációja tipikusan politikai határok (országok, megyék, városok) mentén történik, ugyanakkor a vizsgálat tárgyát képező folyamatok esetleg teljesen más mértéket követnek a térben. Mindezek következményeként az egymással szomszédos térbeli egységekben a vizsgált változó értékei igen közel esnek egymáshoz, ami a statisztikai-ökonometriai elemzésben pozitív térbeli autokorrelációként jelenik meg (példaként lehet megemlíteni azokat az eseteket, amikor a természeti adottságok kihasználására alapuló gazdasági tevékenységek – bányászat, mezőgazdaság stb. – több földrajzi egység területét ölelik fel, vagy amikor az országhatárokon átnyúló együttműködések következményeként a szomszédos, de az országhatárok által elválasztott térbeli egységek között bizonyos hasonlóságok alakulnak ki, ahogyan az például az osztrák-magyar határ mentén megfigyelhető).

A szubsztantív magyarázat szerint a társadalmi interakciók sajátos természetéből következik a térbeli autokorreláció, hiszen a megfigyelési egységek közötti kapcsolatok következtében a vizsgált változó igen hasonló értékei jöhetnek létre egymáshoz közel fekvő lokációkban még akkor is, ha az adatok aggregációja megfelelő. Mindez abból a közismert tényből következik, miszerint a társadalmi interakciók kialakulását a térbeli közelség jelentősen elősegíti. Az ekképpen létrejött autokorreláció erőssége tehát a térben nem konstans, hanem az egységek közötti távolsággal fordított arányban változik. Érvényes tehát az, amit *Tobler* (1979) a geográfia első törvényének nevez, vagyis, hogy „minden mindennel összefügg, de közeli jelenségek sokkal inkább összefüggenek egymással, mint távoli dolgok” (*Tobler*; 1979).

Amint az az előzőekben megállapítást nyert, abban az esetben, ha a térbeli adatokat homogenitás vagy térbeli függetlenség jellemzi, az időbeli adatoknál használt

ökonometriai megoldások használata elfogadható. Amennyiben azonban a térbeli adatok heterogének és (vagy) közöttük autokorreláció tapasztalható, akkor az alkalmazandó regressziós eljárás szempontjából a megfigyelések térbeli pozíciója kiemelt jelentőségre tehet szert, s a térbeli helyzet megfelelő ökonometriai–statisztikai reprezentációja a további elemzés alapfeltételévé válhat. Az adatok térbeli heterogenitása eredményezheti a regressziós egyenletben jelentkező heteroszkedaszticitást. A térbeli heterogenitás kezelése így viszonylag kisebb kihívással jár, hiszen az esetek többségében a heteroszkedaszticitással kapcsolatos problémáknál széles körben alkalmazott ökonometriai módszerek, mint például a véletlen koefficiens modellek (random coefficient models), a hiba komponens modellek (error component models), vagy az adaptív szűrés (adaptive filtering) módszerei jól használhatók.<sup>4</sup> Mindezek mellett a térbeli adatelemzés irodalmában kifejezetten térspecifikus módszereket is találunk. E modellekben a tér reprezentációjára két alapvető megoldás található: a megfigyelési egységeket vagy azok *koordinátái* révén [ahogyan ezzel a módszerrel például a trend felszín analízisben (trend surface analysis) vagy a térbeli paraméter kiterjesztés modelljében (spatial expansion model) találkozhatunk] vagy pedig *dummy változók* felhasználásával [mint például a térbeli variancia-analízisben (spatial analysis of variance) vagy a térben változó regresszió módszerében (spatially switching regression)] építik be az elemzésbe.<sup>5</sup> Míg a térbeli heterogenitás kezelése lényegében nem igényel speciális ökonometriai kezelésmódokat, addig ez a térbeli autokorreláció tárgyalásáról nem mondható el. Mindezek miatt a továbbiakban a tanulmány a térbeli függőség kezelésének módszertanára összpontosít.

A térbeli autokorreláció ökonometriai tárgyalásához olyan térreprezentációra van szükség, amely képes az egységek egymáshoz viszonyított helyzetét megragadni. A megfigyelési egységek szomszédsági fokok szerinti kategorizálása széleskörűen használt kiindulópont a relatív térbeli helyzetek leképezésében. Az elsőfokú szomszédság fogalma az egységek közötti legnagyobb közelség (legkisebb távolság) megragadására szolgál. Egymással közös határral bíró, vagy egymástól egy bizonyos kritikus távolságon belül elhelyezkedő megfigyelési egységeket nevezünk elsőfokú szomszédoknak. A másod- harmad- stb. fokú szomszédság pedig a közös határon, vagy kritikus távolságon alapuló felfogásnak a távolabb fekvő egységekre való kiterjesztéseként adódik (a másodfokú szomszéd az elsőfokú szomszéd szomszédja és így tovább).

A megfigyelési egységek relatív térbeli pozícióinak megragadására szolgálnak az úgynevezett térbeli súlymátrixok (spatial weights matrices). A térbeli súlymátrix ( $\mathbf{W}$ ) dimenziója a megfigyelési egységek számával ( $N$ ) egyenlő. A térstruktúra legegyszerűbb leképezése az úgynevezett bináris szomszédságmátrixokkal (binary contiguity matrices) lehetséges, amikor a mátrix bármely elemének ( $w_{ij}$ ) két lehetséges értéke 1 és 0. Amikor  $i$  és  $j$  egymással (a választott szomszédsági fok szerint) szomszédos, akkor az érték 1 lesz, egyébként pedig 0. Megállapodás szerint a mátrix diagonális elemei nullával egyenlők. A bináris szomszédságmátrixok szimmetrikusak.

<sup>4</sup> A heteroszkedasztikus reziduummal összefüggő ökonometriai problémák kezelésére általánosan alkalmazott módszerek színvonalas áttekintését magyar nyelven lásd *Kőrösi, Máttyás és Székely* (1990) könyvében, valamint *Pintér* (1991) tanulmányában.

<sup>5</sup> Bővebben e módszerekről lásd például *Anselin* (1998a), *Casetti* (1997), *Griffith* (1988), *Haining* (1990) vagy *Kristensen* (1996).

A következő mátrix egy 4x4-es bináris szomszédságmátrix példája:

$$\mathbf{W} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

A mátrixok másik osztályába tartoznak az úgynevezett inverz távolságalapú súlymátrixok (inverse distance weights matrices). Ekkor a mátrix egy-egy eleme az  $i$  és  $j$  lokációk közötti távolság reciprokának valamely fokú hatványával egyenlő.<sup>6</sup> A következő, 4x4-es mátrix erre szolgál példaként, ahol  $d_{ij}$  az  $i$ -edik és a  $j$ -edik megfigyelési egység közötti távolságot reprezentálja.

$$\mathbf{W} = \begin{pmatrix} 0 & \frac{1}{(d_{1,2})^2} & \frac{1}{(d_{1,3})^2} & \frac{1}{(d_{1,4})^2} \\ \frac{1}{(d_{2,1})^2} & 0 & \frac{1}{(d_{2,3})^2} & \frac{1}{(d_{2,4})^2} \\ \frac{1}{(d_{3,1})^2} & \frac{1}{(d_{3,2})^2} & 0 & \frac{1}{(d_{3,4})^2} \\ \frac{1}{(d_{4,1})^2} & \frac{1}{(d_{4,2})^2} & \frac{1}{(d_{4,3})^2} & 0 \end{pmatrix}.$$

A súlymátrixok minden egyes elemének a saját sorában elhelyezkedő elemek összegével való osztása révén kapjuk az úgynevezett sorstandardizált súlymátrixokat (row-standardized weights matrices). Értelemszerűen a sorstandardizált súlymátrixok minden egyes eleme nulla és egy közötti értéket vehet fel s az elemek összege minden egyes sorban 1-gyel egyenlő. A sorstandardizált mátrixok használata két okkal magyarázható: egyrészt az általuk kapott eredmények könnyebben értelmezhetők, másrészt a maximum likelihood becsléssel összefüggő számítások csak ezek alkalmazásával végezhetők el, ahogyan az a későbbiekben bemutatásra kerül. A térbeli autokorreláció statisztikai-ökonometriai kezelésének mélyebb megvilágítása céljából a következő fejezet a térbeli függőség tesztelésének módszertanába nyújt betekintést.

## A TÉRBELI AUTOKORRELÁCIÓ STATISZTIKAI TESZTELÉSE

A térbeli autokorreláció<sup>7</sup> jelenlétére az egymáshoz igen hasonló értékek térbeli csoportosulásából (pozitív térbeli autokorreláció), vagy az egymással szomszédos megfigyelési egységeknél igen eltérő értékek jelentkezéséből (negatív térbeli autokorreláció) következtethetünk. A pusztán vizuális benyomás azonban nem helyettesítheti annak statisztikai

<sup>6</sup> A térbeli súlymátrixokról jóval részletesebben lásd a következő munkákat: *Anselin* (1988a), *Cliff* és *Ord* (1981) és *Upton* és *Fingleton* (1985).

<sup>7</sup> A térbeli autokorreláció statisztikai vonatkozásairól magyar nyelven lásd még *Nemes Nagy* (1998).

tikai tesztelését, hogy vajon az adatok adott térbeli konstellációja pusztán a véletlen eredménye, vagy pedig térbeli autokorreláció jelenlétének következménye. A térbeli autokorreláció tesztelésére két típusú próbacsoport terjedt el a gyakorlatban: az egyik csoportba az úgynevezett globális próbafüggvények, míg a másikba a lokális statisztikák tartoznak. A globális tesztek az adott változó térbeli autokorrelációra való hajlamát tesztelik, a teljes adatsor figyelembevételével, míg a lokális próbák az egyes megfigyelési egységek szintjén vizsgálják a helyi (pozitív, vagy negatív) autokorreláció jelenlétét.

A továbbiakban a globális próbafüggvények közül a Moran-féle  $I$ -próba (Moran's  $I$  statistic) (Moran; 1948, Cliff-Ord; 1981), míg a lokális tesztek közül az Anselin-féle Lokális Moran-próba (Local Moran) (Anselin; 1995) kerül bemutatásra.<sup>8</sup>

A Moran-féle  $I$ -próba próbafüggvénye:

$$I = [N/S_0] [\sum_{i,j} w_{ij} (x_i - \mu)(x_j - \mu) / \sum_i (x_i - \mu)^2], \quad /1/$$

ahol  $N$  a megfigyelési egységek száma,  $x_i$  és  $x_j$  az  $x$  változó két pontban mért értékei,  $\mu$  ezek várható értéke,  $w_{ij}$  a térbeli súlymátrix egy eleme,  $S_0$  pedig normalizáló faktor:

$$S_0 = \sum_{i,j} w_{ij}, \quad /2/$$

vagyis a súlyok összege. Ha  $\mathbf{W}$  sorstandardizált súlymátrix (ahogyan az általában javasolt), akkor  $S_0 = N$ , és a próbafüggvény a következő egyszerűbb formát ölti:

$$I^* = \sum_{i,j} w_{ij} (x_i - \mu)(x_j - \mu) / \sum_i (x_i - \mu)^2, \quad /3/$$

ami bár hasonló, de mégsem teljesen azonos a korrelációs együttathóval.  $I^*$  várható értéke ugyanis  $[-1/(N-1)]$ -el egyenlő, ami a mintaelemszám növekedésével a nullához közelít (Anselin; 1992).  $I^*$ -nak a várható értéknél nagyobb értékei pozitív, míg annál kisebb értékei negatív térbeli autokorrelációra utalnak.

A globális térbeli autokorreláció tesztelésére használatos bármely próba (melyek közül most a Moran-féle  $I$ -próba került bemutatásra) ugyanarra az alapelvre épül, mely az úgynevezett Gamma-indexben (Hubert-Constanzo-Gale; 1985) nyer általános formát:

$$\Gamma = \sum_{i,j} w_{ij} c_{ij} . \quad /4/$$

A Gamma-index jelöléseit alkalmazva, a Moran-féle  $I$ -próba függvényében a  $c_{ij}$  a következőképpen határozható meg:

$$c_{ij} = (x_i - \mu)(x_j - \mu) / \sum_i (x_i - \mu)^2. \quad /5/$$

A Gamma-indexre alapuló globális próbák a következő elven alapulnak. Mivel a térbeli autokorreláció vagy hasonló értékek (pozitív térbeli autokorreláció), vagy pedig nagyon különböző értékek (negatív térbeli autokorreláció) térbeni csoportosulását jelenti, ezért a próbák a megfigyelési egységekben mért értékek között fennálló tendenciát (vagyis, hogy azok

<sup>8</sup> A térbeli autokorreláció egyéb globális és lokális próbafüggvényeiről lásd részletesebben például Anselin (1995), Getis és Ord (1992) vagy Cliff-Ord (1981).



tendenciaszerűen hasonló, vagy pedig különbözők-e egymástól) a megfigyelési egységek térbeli helyzeteiben észlelt tendenciákkal (vagyis, hogy a hasonló, vagy különböző értékek általában térben „közel”, vagy pedig „távol” helyezkednek-e el egymástól) vetik össze. Ha az  $I$  egyenletben  $c_{ij} > 0$  (vagyis mindkét lokációban a mért érték nagyobb, vagy kisebb mint a várható érték), és  $w_{ij}$  értéke is nullától különbözik (vagyis  $i$  és  $j$  szomszédai egymásnak), akkor hasonló értékek találhatók egymással szomszédos lokációkban, ami pozitív térbeli autokorrelációt jelent. Abban az esetben, ha különböző értékek egybeesése tapasztalható a térben, akkor ugyanazon  $i$  és  $j$  lokációkra a  $c_{ij}$  negatív értéket, míg a  $w_{ij}$  nullától különböző értéket mutat, ami a negatív térbeli autokorreláció jelenlétére utal.

A Moran-féle  $I$ -próba szignifikanciájának, vagyis annak eldöntésére, hogy vajon egy változó értékeinek adott térbeli eloszlása térbeli autokorreláció jelenlétének eredménye-e vagy sem, két alapvető módszer használatos. Mindkét módszer az értékek adott térbeli eloszlását az általában véletlenszerűen kialakuló térbeli eloszlásokhoz viszonyítja, s ha az adott térbeli eloszlás extrémnek minősül a véletlenszerűen előfordulóhoz képest, akkor ezt szignifikáns térbeli autokorrelációnak (azaz az értékeknek a véletlenszerűen előfordulótól eltérő mértékű térbeli csoportosulásának) minősíti. Természetesen az, hogy mi minősül extrémnek, részben a változónak a null-hipotézis (azaz a térbeli autokorreláció nélküli eset feltételezése) melletti valószínűségeloszlásától, részben a választott szignifikanciaszinttől függ.

A leggyakrabban használatos eljárás az értékek normális eloszlásának feltételezésére épül. Ekkor a Moran-féle  $I$ -próba szignifikanciájának megállapítása a standardizált normális eloszlású valószínűségi változó értékén alapul, melyet a szokásos módon képezünk:

$$Z_I = [I - E(I)] / SD(I), \quad /6/$$

ahol  $E(I)$  az  $I$  várható értéke,  $SD(I)$  pedig az  $I$  szórása. Az  $I$  szignifikanciájának megállapítása pedig az általában szokásos módon történik.

Az értékek normális eloszlásának feltevése azonban igen sok esetben túlságosan erős feltételezés. Alternatívaként az úgynevezett permutációs módszer használata ajánlott. A módszer null-hipotézise szerint annak valószínűsége, hogy a mért értékek milyen permutációban jelennek meg az egyes lokációk között, bármely permutációra vonatkozólag ugyanaz. E feltételezés alapján a változó mért értékeinek a lokációk közötti véletlenszerű permutációihoz kiszámítjuk az  $I$  megfelelő értékeit, melynek eredményeként az  $I$  értékeinek úgynevezett mesterséges referenciaeloszlása hozható létre. Az extremitás fokának megállapításakor a tényleges  $I$  értéket viszonyítjuk a véletlen permutációk eredményeként számított  $I$  értékek eloszlásához. Egy egyszerű hüvelykujjszabály szerint a döntés alapja az úgynevezett pszeudo-szignifikancia szint. Ezt a következőképpen számítjuk:  $(T+1)/(M+1)$ , ahol  $T$  azoknak az eseteknek a száma, amikor a referenciaeloszlás által adott értékek nagyobbak vagy egyenlők, mint a mért  $I$  érték, míg  $M$  a permutációk számát jelöli. A pszeudo-szignifikancia szint alacsony értéke annak jele, hogy az értékek adott térbeli eloszlása extrém esetnek minősül ahhoz képest, ami a null-hipotézis alapján elvárható lenne, ami (az  $I$  előjelétől függően) pozitív, vagy negatív térbeli autokorreláció jelenlétét igazolja.

A térbeli autokorreláció globális próbái kiválóan alkalmazhatók az adatok térbeli csoportosulásra való hajlamának megállapítására. Sok esetben viszont nemcsak az általános, de a helyi tendenciák ismerete is szükséges. Előfordulhat, hogy az adatokat általában nem jellemző térbeli autokorreláció, mégis található néhány határozott térbeli csoportosulás (klaszter).

Az egyes helyi csoportosulások szignifikanciájának megállapítására a térbeli autokorreláció lokális próbafüggvényeit alkalmazhatjuk. Ezek közül a következőkben az Anselin-féle Lokális Moran-próba kerül bemutatásra. A próba függvénye:

$$I_i = z_i \sum_j w_{ij} z_j, \quad /7/$$

ahol  $I_i$  az  $i$ -edik megfigyelési egységre számított Lokális Moran próbafüggvény értéke,  $z_i = x_i - \mu$ , vagyis  $z_i$  a változó értékének és a várható értéknek a különbsége az  $i$ -edik lokációban,  $z_j$  pedig ugyanez a  $j$ -edik lokációra. Mivel  $\mathbf{W}$  sorstandardizált, ezért a  $\sum_j w_{ij} z_j$  tag az  $i$ -edik megfigyelési egységhez tartozó szomszédsági értékek súlyozott átlagát jelenti. Ha  $z_i$  és  $\sum_j w_{ij} z_j$  értékei egyként pozitívak (negatívak), akkor  $I_i$  értéke pozitív autokorrelációt mutat, ami magas (alacsony) értékek térbeli koncentrációját jelenti az  $i$ -edik megfigyelési egység környezetében. Az  $I_i$  szignifikanciájának meghatározásához a globális Moran  $I$ -próbafüggvényében is alkalmazott permutációs módszer használatos.

#### TÉRBELI AUTOKORRELÁCIÓ A LINEÁRIS TÉRÖKONOMETRIAI MODELLEKBEN

Amint a következőkből is látható, abban az esetben, ha az adatokat térbeli autokorreláció jellemzi, a hagyományosan használt becslési módszerek nem alkalmazhatók. A térbeli autokorreláció ökonometriai modellezésének két legelterjedtebb módszere a térbeli késleltetés modellje (spatial lag model), illetve a térbeli hiba autokorreláció modellje (spatial error model).

##### *A térbeli késleltetés modellje*

A súlymátrixoknak a térbeli regressziós modellekben való alkalmazásával egy változó bizonyos pontban mért értékét ugyanezen változónak a tér más pontjain mért értékeivel hozzuk összefüggésbe. Az idősoros elemzés mintájára térbeli adatokra is értelmezhető a késleltetés koncepciója, azzal a különbséggel, hogy itt a késleltetés nem az időben, hanem a térben való „elcsúszásként” értelmezett.

Az idősoros elemzés során a késleltetett érték a késleltetési operátor révén kerül kifejezésre, vagyis például

$$L^k y_t = y_{t-k}, \quad /8/$$

ahol  $L^k y_t$  az  $y_t$   $k$  időszakkal korábbi értékét jelöli. Térbeli adatok esetén sajnos nem ilyen egyértelmű a helyzet, ami annak következménye, hogy a késleltetés bármely szomszédsági fokon belül többirányú is lehet. Hacsak a mintaelemszám nem kifejezetten magas, a megoldás a térben késleltetett változók nagy számát eredményezheti, ami miatt a szabadfok olyan kicsire zsugorodhat, ami akár a regressziós egyenlet becslését is lehetetlenné teheti. A probléma megoldására a térbeli késleltetés során az adott szomszédsági fokba tartozó megfigyelési egységekben mért értékek súlyozott átlagát vesszük figyelembe, vagyis

$$L^s x_i = \sum_j w_{ij} x_j, \quad /9/$$

ahol  $L^s$  az  $s$ -edik fokú szomszédságnak megfelelő térbeli késleltetési operátor,  $x_j$  az  $x$ -nek a  $j$ -edik megfigyelési egységben mért értéke, ahol  $j$  az  $s$ -edik szomszédsági foknak megfelelő eleme a térnek,  $w_{ij}$  pedig a sorstandardizált térbeli súlymátrix megfelelő eleme.<sup>9</sup>

A térbeli késleltetés modelljének általános formája:

$$\mathbf{y}_{(N \times 1)} = \rho \mathbf{W}_{(N \times N)} \mathbf{y}_{(N \times 1)} + \mathbf{X}_{(N \times K)} \boldsymbol{\beta}_{(K \times 1)} + \boldsymbol{\varepsilon}_{(N \times 1)}, \quad /10/$$

ahol  $\mathbf{y}$  az eredményváltozó értékeinek vektora,  $\mathbf{W}$  sorstandardizált súlymátrix,  $\mathbf{W}\mathbf{y}$  az eredményváltozó térben késleltetett értékeinek vektora,  $\mathbf{X}$  az exogén változók mátrixa,  $\rho$  a térbeli autoregressziós paraméter,  $\boldsymbol{\beta}$  az exogén változók paramétervektora,  $\boldsymbol{\varepsilon}$  pedig az egymástól független és azonos valószínűségeloszlású hibatagok vektora. A hibatagok várható értéke  $\mathbf{0}$ , szórásnégyzete pedig  $\sigma^2$ .

Míg az időbeli adatokra épülő késleltetési modell becslésére (abban az esetben, ha hiba-autokorreláció nem tapasztalható) alkalmazható a Legkisebb Négyzetek Módszere (Ordinary Least Squares – OLS) (lásd például *Greene*; 1993), mindez nem igaz a térbeli késleltetés modelljének becslésére. Ennek oka a térben késleltetett függő változó endogenitása (hiszen a térbeli autokorreláció kétirányúsága következtében  $\mathbf{W}\mathbf{y}$  nemcsak meghatározza az  $\mathbf{y}$ -t, de maga is meghatározott az  $\mathbf{y}$  által). Következésképpen a térben késleltetett függő változó korrelál a hibataggal, s így az OLS-becslőfüggvény torzított és inkonzisztens lesz (*Anselin*; 1988a). Az OLS helyett leggyakrabban használt két becslési eljárás a maximum likelihood módszer, illetve az instrumentális változók módszere.

A maximum likelihood becslés<sup>10</sup> alkalmazásának alapfeltétele a reziduális változó normális eloszlása. A becslési eljárás során azokat a paramétereket keressük, melyek mellett a függőváltozó megfigyelt értékei együttes előfordulásának valószínűsége maximális. A függőváltozó likelihood függvénye a normális eloszlást követő reziduális változó együttes sűrűségfüggvényének transzformációjaként adódik. A likelihood függvényt a reziduumok együttes sűrűségfüggvényének és a Jakobi-mátrix determinánsának a szorzataként kapjuk. A Jakobi mátrix a reziduális változónak a függőváltozó szerinti elsőfokú deriváltjait tartalmazó mátrix. A térbeli késleltetés modelljében a Jakobi mátrix determinánsának formája:  $|\mathbf{I} - \rho \mathbf{W}|$ .

Így a logaritmus likelihood függvény a következő formát ölti:

$$L = \ln |\mathbf{I} - \rho \mathbf{W}| - N/2 \ln (2\pi) - N/2 \ln (\sigma^2) - (\mathbf{y} - \rho \mathbf{W}\mathbf{y} - \mathbf{X}\boldsymbol{\beta})' (\mathbf{y} - \rho \mathbf{W}\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) / 2 \sigma^2. \quad /11/$$

A logaritmus likelihood függvény  $\boldsymbol{\beta}$ ,  $\rho$ , és  $\sigma^2$  szerinti maximalizálása eredményezi a paraméterek azon értékeit, melyek mellett a függőváltozó adott mintájának előfordulási valószínűsége a legnagyobb. Mivel e probléma a térbeli késleltetés modelljére analitikusan nem megoldható, ezért a gyakorlatban numerikusan történik a paraméterek meghatározása.

Az egyébként igen bonyolult számításokat két tény jelentősen könnyíti. Egyrészt, mivel a  $\boldsymbol{\beta}$  és  $\sigma^2$  kifejezhető  $\rho$  függvényeként ezek helyettesíthetők az előbbi logaritmus likelihood

<sup>9</sup> A térbeli késleltetés koncepcójának igen részletes magyarázatára lásd *Anselin* (1988a).

<sup>10</sup> A maximum likelihoodnak a térbeli késleltetés modelljében történő alkalmazásáról igen részletes elemzés olvasható a következő munkákban: *Anselin* (1988a), *Cliff és Ord* (1981) vagy *Ord* (1975).

függvénybe, ami az úgynevezett koncentrált likelihood függvényt eredményezi. Ugyanakkor a Jakobi-mátrix determinánsa az előzőnél lényegesen egyszerűbb formában is kifejezhető (Ord; 1975). Eszerint  $\ln |\mathbf{I} - \rho \mathbf{W}| = \sum_i \ln(1 - \rho \omega_i)$ , ahol  $\omega_i$  a  $\mathbf{W}$  saját értékeit jelöli. Ennek az eredménynek igen nagy a gyakorlati jelentősége, hiszen a Jakobi-mátrix determinánsának ezen formája lényegesen megkönnyíti a paraméterek kiszámítását.

A koncentrált likelihood függvény az előzők alapján a következő formát ölti:

$$L_c = \sum_i \ln(1 - \rho \omega_i) - (N/2) \ln[(\mathbf{e}_0 - \rho \mathbf{e}_L)' (\mathbf{e}_0 - \rho \mathbf{e}_L) / N], \quad /12/$$

ahol  $\mathbf{e}_0$  reziduum vektor abban a regressziós egyenletben, ahol  $\mathbf{y}$  a bal oldali és  $\mathbf{x}$  a jobb oldali változó,  $\mathbf{e}_L$  pedig szintén reziduum vektor akkor, amikor  $\mathbf{W}\mathbf{y}$  szerepel bal oldali és  $\mathbf{x}$  jobb oldali változóként. A  $\rho$  elfogadható értékei az  $1/\omega_{\max}$ -tól az  $1/\omega_{\min}$ -ig terjedő sávban található, ahol  $\omega_{\min}$  a  $\mathbf{W}$  legkisebb,  $\omega_{\max}$  pedig a  $\mathbf{W}$  legnagyobb saját értéke. A  $\rho$  azon értékének meghatározása, melynél a korábbi koncentrált likelihood függvény a maximális értéket veszi fel, numerikusan történik. Ezek után behelyettesítéssel adódik a  $\beta$  és  $\sigma^2$  megfelelő értéke. A becslés további részleteiről lásd *Anselin* (1988a).

Ha a reziduum nem normális eloszlást követ, akkor a leggyakrabban használt alternatív becslési eljárás a kétfokozatú legkisebb négyzetek (2SLS) módszere. A 2SLS alkalmazását a térben késleltetett függőváltozó endogenitása indokolja. Az endogenitásból következően e változó korrelál a reziduummal. A 2SLS során instrumentális változók felhasználásával az endogén változót annak becslött értékével helyettesítjük, mely már mentes a reziduummal való autokorrelációtól. Instrumentális változóként az egyenlet első fokon késleltetett magyarázó változóinak használata javasolt az irodalomban ( $\mathbf{W}\mathbf{X}$ ).<sup>11</sup>

#### *A térbeli hiba autokorreláció modellje*

A térbeli autokorreláció ökonometriai modellezésére különböző lehetőségek adódnak. A térbeli késleltetés előzőkben vázolt modelljében a térbeli függőség közvetlen módon kerül modellezésre. Ez a közelítésmód más néven szubsztantív térbeli függőségmodellezésként is szerepel az irodalomban,<sup>12</sup> utalva arra, hogy a térben késleltetett független változó becslött paraméterének értéke és annak szignifikanciája explicit módon tudósít arról, hogy a független változó adott értékei kialakulásában a változónak a tér más pontjain mért értékei milyen szerepet játszanak.

Alternatív lehetőségként kínálkozik a térbeli függőség olyan megközelítése, amikor a regressziós egyenlet a magyarázó változóknak és a független változónak a térbeli autokorreláció hatásaitól „megtisztított” viszonyát tárja elénk. Ekkor a térbeli autokorreláció zavaró, vagyis kiiktatandó tényezőként szerepel a modellben. A térbeli hiba autokorreláció modellje tehát a hibatagok között jelentkező térbeli autokorreláció korrekcióját szolgálja.

A térbeli hiba autokorreláció modelljének általános formája:

$$\mathbf{y}_{(N \times 1)} = \mathbf{X} \beta_{(K \times 1)} + \boldsymbol{\varepsilon}, \quad /13/$$

<sup>11</sup> A térbeli késleltetés modelljének instrumentális változók módszerével való becsléséről lásd még *Land és Deane* (1992), *Kelejian és Robinson* (1993), *Kelejian és Prucha* (1998). Az instrumentális változók módszerének általánosabb vonatkozásairól pedig magyar nyelven bővebben *Kőrösi, Mátyás és Székely* (1990) kínál áttekintést.

<sup>12</sup> Lásd például *Anselin* (2001).

és

$$\boldsymbol{\varepsilon}_{(N \times 1)} = \lambda \mathbf{W} \boldsymbol{\varepsilon} + \boldsymbol{\zeta}, \quad /14/$$

ahol  $\mathbf{y}$  az eredményváltozó értékeinek vektora,  $\mathbf{X}$  az exogén változók mátrixa,  $\boldsymbol{\beta}$  az exogén változók paramétervektora,  $\boldsymbol{\varepsilon}$  az autoregresszív hibatagok vektora,  $\mathbf{W}$  sorstandardizált súlymátrix,  $\lambda$  az autoregresszív hibatagok térben késleltetett értékeinek paramétere,  $\boldsymbol{\zeta}$  pedig az egymástól független és azonos valószínűségeloszlású hibatagok vektora. A független hibatagok várható értéke  $\mathbf{0}$ , szórásnégyzete pedig  $\sigma^2$ .

Az idősoros esethez hasonlóan, az OLS becslőfüggvény torzítatlan, de nem hatásos a térbeli hiba autokorreláció esetében sem.<sup>13</sup> Szemben viszont az idősoros esettel, az ott igen gyakran használt általánosított legkisebb négyzetek módszere (GLS) a térbeli adatokra nem alkalmazható (Anselin; 1988a). Ennek oka a térbeli autokorreláció szimultán jellege, ami miatt a térbeli hiba autokorreláció esetén a maximum likelihood a leggyakrabban használatos becslési eljárás.

A térbeli hiba autokorreláció esetén a logaritmusos likelihood függvény a következő alakot ölti:

$$L = \ln |\mathbf{I} - \lambda \mathbf{W}| - N/2 \ln(2\pi) - N/2 \ln(\sigma^2) - (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{I} - \lambda \mathbf{W})'(\mathbf{I} - \lambda \mathbf{W})(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})/2\sigma^2 \quad /15/$$

a korábban már ismertetett jelöléseket alkalmazva. A likelihood függvényt  $\lambda$ ,  $\boldsymbol{\beta}$  és  $\sigma^2$  szerint maximalizálva kapjuk a paraméterek maximum likelihood becslését.<sup>14</sup> A térbeli hiba autokorreláció modellje alternatív becslési eljárásainak kifejlesztése – mint például az általánosított momentum módszer (GMM) alkalmazási lehetőségeinek kidolgozása – napjaink térökonometriai kutatásainak egyik igen aktív területe.<sup>15</sup>

### A TÉRBELI AUTOKORRELÁCIÓ TESZTELESE ÖKONOMETRIAI MODELLEKBEN

A legmegfelelőbb térökonometriai modellspecifikáció meghatározását számos teszt segíti. A diagnosztikák első csoportjába azok a próbák tartoznak, amelyek a térbeli hiba autokorreláció vagy a térbeli késleltetés modelljei közötti döntést teszik lehetővé. Ezek a próbák a térbeli autokorrelációnak az OLS-módszerrel becsült regressziós egyenletben való jelenlétét vizsgálják. A második csoportban pedig olyan diagnosztikák találhatók, melyek az alkalmazott térökonometriai modell megfelelőségét tesztelik abból a szempontból, hogy vajon a szóban forgó modell a térbeli autokorreláció problémáját megnyugtatóan kezeli-e vagy valamilyen formában a függőség még mindig jelen van a regressziós egyenletben. A

<sup>13</sup> A különböző becslési eljárásokról, illetve az esztimátorok tulajdonságairól részletesebben lásd Anselin (1988a), Cliff és Ord (1981) és Ord (1975).

<sup>14</sup> A térbeli késleltetés modelljéhez hasonlóan  $\boldsymbol{\beta}$  és  $\sigma^2$  a  $\lambda$  függvényeként adódik, ami a következő koncentrált likelihood függvényt eredményezi:  $L_c = \sum_i \ln(1 - \lambda \omega_i) - (N/2) \ln[(1/N) \mathbf{e}'(\mathbf{I} - \lambda \mathbf{W})'(\mathbf{I} - \lambda \mathbf{W})\mathbf{e}]$ , ahol  $\mathbf{e}$  a GLS becslés reziduuma vektora, míg a többi jelölés a fentiekkel megegyezik. A  $\lambda$  elfogadható értékei az  $1/\omega_{\max}$ -tól az  $1/\omega_{\min}$ -ig terjedő sávban találhatóak, ahol  $\omega_{\min}$  a  $\mathbf{W}$  legkisebb,  $\omega_{\max}$  pedig a  $\mathbf{W}$  legnagyobb saját értéke. A  $\lambda$  azon értékének meghatározása, melynél a fenti koncentrált likelihood függvény a maximális értéket veszi fel, numerikusan történik. Ezek után behelyettesítéssel adódik a  $\boldsymbol{\beta}$  és  $\sigma^2$  megfelelő értéke. A becslés további részleteiről lásd Anselin (1988a).

<sup>15</sup> Bővebben lásd Conley (1996), Kelejian és Prucha (1999).

következőkben a legfontosabb próbák bemutatása következik.<sup>16</sup> A próbafüggvények összefoglaló bemutatását lásd a táblában.

*Az ökonometriai modellekben leggyakrabban használt térbeli autokorreláció-tesztek*

A próba neve	Próbafüggvény	Valószínűség-eloszlás	Forrás
MORAN	$\mathbf{e}'\mathbf{W}\mathbf{e}/\mathbf{e}'\mathbf{e}$	$N(0,1)$	<i>Cliff-Ord</i> (1981)
LM-ERR	$(\mathbf{e}'\mathbf{W}\mathbf{e}/s^2)/T$	$\chi^2(1)$	<i>Burridge</i> (1980)
LM-ERRLAG	$(\mathbf{e}'\mathbf{W}\mathbf{e}/s^2) / [\text{tr}(\mathbf{W}'\mathbf{W} + \mathbf{W}^2) - \text{tr}(\mathbf{W}'\mathbf{W} + \mathbf{W}^2)\mathbf{A}^{-1}\text{var}(\rho)]$	$\chi^2(1)$	<i>Anselin</i> (1988a)
LM-LAG	$(\mathbf{e}'\mathbf{W}\mathbf{y}/s^2) / (\text{RJ}_{\rho-\beta})$	$\chi^2(1)$	<i>Anselin</i> (1988b)
LM-LAGERR	$(\mathbf{e}'\mathbf{B}'\mathbf{B}\mathbf{W}\mathbf{y})^2 / (H_{\rho} - H_{0\rho} \text{Var}(\boldsymbol{\theta})\mathbf{H}'_{0\rho})$	$\chi^2(1)$	<i>Anselin-Bera-Florax-Yoon</i> (1996)

A térbeli hiba autokorrelációnak az OLS modellben való tesztelésére széleskörűen alkalmazott módszer a Moran-féle *I*-próba következő formája:

$$I = \mathbf{e}'\mathbf{W}\mathbf{e}/\mathbf{e}'\mathbf{e}, \quad /16/$$

ahol  $\mathbf{e}$  az OLS-reziduumok vektora,  $\mathbf{W}$  pedig sorstandardizált ( $N \times N$ ) méretű súlymátrix.

A statisztikai hipotézisvizsgálat a normális valószínűségeloszlás bázisán történik.<sup>17</sup> Ahogyan azt *Anselin* és *Rey* (1991) kimutatta, a Moran-féle *I*-próba ismertett formája nem megbízható olyan esetekben, ha egyéb specifikációs problémák (például heteroszkedaszticitás) is jellemzik az empirikus modellt. Így a Moran-féle *I*-próba akkor is jelezhet térbeli autokorrelációt, amikor a gond egyéb okokból származik. Mindezek alapján a Moran *I*-próbát egyfajta általános specifikációs probléma-indikátorként ajánlatos lefuttatni, de a hiba-autokorreláció specifikus tesztelésére egyéb próbák elvégzése ajánlott. A Moran-féle *I*-próba kiváltására számos asszimptotikus próba található az irodalomban. A térbeli hiba autokorreláció tesztelésére *Burridge* (1980) fejlesztette ki a következő Lagrange Multiplikátor (LM) próbát:

$$\text{LM-ERR} = (\mathbf{e}'\mathbf{W}\mathbf{e}/s^2)/T, \quad /17/$$

ahol  $\mathbf{e}$  az OLS-reziduumok vektora,  $\mathbf{W}$  térbeli súlymátrix,  $s^2 = \mathbf{e}'\mathbf{e}/N$  és  $T = \text{tr}(\mathbf{W}'\mathbf{W} + \mathbf{W}^2)$ , ahol  $\text{tr}$  a mátrix nyoma (trace) mátrix műveletet jelzi. A próbafüggvény  $\chi^2$  eloszlást követ, egyes szabadságfokkal.

A térben késleltetett függő változó formájában megjelenő térbeli autokorreláció jelenlétére a következő Lagrange Multiplikátor teszt alkalmazása a legelterjedtebb (*Anselin*; 1988b):

$$\text{LM-LAG} = (\mathbf{e}'\mathbf{W}\mathbf{y}/s^2) / (\text{RJ}_{\rho-\beta}), \quad /18/$$

ahol  $\mathbf{e}$  az OLS-reziduumok vektora,  $\mathbf{y}$  a regresszió függő változójának vektora,

$$\text{RJ}_{\rho-\beta} = [T + (\mathbf{W}\mathbf{X}\boldsymbol{\beta})' \mathbf{M}(\mathbf{W}\mathbf{X}\boldsymbol{\beta}) / s^2], \quad /19/$$

<sup>16</sup> A diagnosztikák jóval részletesebb ismertetéséhez lásd *Anselin* (1988a, 1988b) valamint *Anselin* és *Florax* (1995).

<sup>17</sup> A hipotézisvizsgálattal kapcsolatban további részletekkel *Cliff* és *Ord* (1981) munkája szolgál.

ahol  $\mathbf{WX}\boldsymbol{\beta}$  az OLS-modell-előjelzések térben késleltetett értéke, valamint

$$\mathbf{M} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}', \quad /20/$$

a projekciós mátrix. A próba  $\chi^2$  eloszlást követ, egyes szabadságfokkal.

A térbeli hiba autokorrelációnak a térbeli késleltetés modelljében való jelenlétére a következő Lagrange Multiplikátor teszt (*Anselin*; 1988b) használatos:

$$\text{LM-ERRLAG} = (\mathbf{e}'\mathbf{W}\mathbf{e}/s^2) / [\text{tr}(\mathbf{W}'\mathbf{W} + \mathbf{W}^2) - \text{tr}(\mathbf{W}'\mathbf{W} + \mathbf{W}^2)\mathbf{A}^{-1}\text{var}(\rho)], \quad /21/$$

ahol  $\mathbf{e}$  a térbeli késleltetés modellje maximum likelihood reziduumainak vektora,

$$\mathbf{A} = (\mathbf{I} - \rho\mathbf{W}), \quad /22/$$

ahol  $\rho$  a térbeli autoregressziós paraméter, melynek becsült aszimptotikus szórásnégyzete  $\text{var}(\rho)$ . A próbafüggvény  $\chi^2$  eloszlást követ, egyes szabadságfokkal. Amennyiben az LM-ERRLAG teszt szignifikáns értéket mutat, akkor a szóban forgó késleltetéses modell nem használható. Megoldásként vagy a regressziós egyenlet szerkezetében való változtatás vagy pedig alternatív súlymátrixok használata javasolt.

Az LM-LAGERR próba a térben késleltetett függő változó formájában megjelenő térbeni dependencia jelenlétét teszteli a hiba autokorreláció modelljében (*Anselin–Bera–Florax–Yoon*; 1996):

$$\text{LM-LAGERR} = (\mathbf{e}_{(N \times 1)}' \mathbf{B}' \mathbf{B} \mathbf{W} \mathbf{y})^2 / (H_p - \mathbf{H}_{\theta\rho(1 \times [K+2])} \mathbf{Var}(\boldsymbol{\theta}) \mathbf{H}'_{\theta\rho(1 \times [K+2])}), \quad /23/$$

ahol  $\mathbf{e}$  a térbeli hiba autokorreláció modellje maximum likelihood reziduumainak vektora,  $\mathbf{y}$  a függő változó vektora,

$$\mathbf{B} = \mathbf{I} - \lambda\mathbf{W}, \quad /24/$$

$$\boldsymbol{\theta}_{([K+2] \times 1)'} = [\boldsymbol{\beta}_{(K \times 1)'}, \lambda, s^2], \quad /25/$$

$\mathbf{Var}(\boldsymbol{\theta})$  pedig a  $\boldsymbol{\theta}$  paraméter vektor becsült szórásnégyzete:

$$H_p = \text{tr}\mathbf{W}^2 + \text{tr}(\mathbf{B}\mathbf{W}\mathbf{B}^{-1})'(\mathbf{B}\mathbf{W}\mathbf{B}^{-1}) + (1/s^2)(\mathbf{B}\mathbf{W}\mathbf{X}\boldsymbol{\beta})'(\mathbf{B}\mathbf{W}\mathbf{X}\boldsymbol{\beta}), \quad /26/$$

$$\mathbf{H}'_{\theta\rho(1 \times [K+2])} = [\mathbf{A}_{1(K \times 1)}, \mathbf{A}_2, 0]', \quad /27/$$

ahol:

$$\mathbf{A}_{1(K \times 1)} = (1/s^2)(\mathbf{B}\mathbf{X})'\mathbf{B}\mathbf{W}\mathbf{X}\boldsymbol{\beta}, \quad /28/$$

és

$$\mathbf{A}_2 = \text{tr}(\mathbf{W}\mathbf{B}^{-1})'\mathbf{B}\mathbf{W}\mathbf{B}^{-1} + \text{tr}\mathbf{W}^2\mathbf{B}^{-1}. \quad /29/$$

A próbafüggvény  $\chi^2$  eloszlást követ, egyes szabadságfokkal. Amennyiben az LM-LAGERR teszt szignifikáns értéket mutat, akkor a szóban forgó térbeli hiba autokorreláció modell nem használható. Megoldásként vagy a regressziós egyenlet szerkezetében való változtatás vagy pedig alternatív súlymátrixok használata javasolt.

A vizsgált jelenség természetének leginkább megfelelő térbeli súlymátrix kiválasztása minden térökonometriai elemzés kritikus pontja. Habár léteznek tesztek, melyek a mátrixválasztáshoz segítséggel szolgálhatnak (lásd *Anselin*; 1984), általánosan alkalmazott egyértelmű megoldás a problémára még nem született. A térbeli autokorreláció tesztelése során ezért több, a teret különbözőképpen strukturáló súlymátrixot alkalmazunk. Amennyiben specifikációs probléma merül fel, mindig az a mátrix választandó, amelyiknél a térbeli autokorreláció problémája a leghatározottabban jelentkezik.

\*

A közgazdaságtan fokozódó érdeklődése a gazdasági jelenségek térbeli vonatkozásai iránt előtérbe helyezi a térbeli modellek ökonometriai tesztelése során megbízhatóan alkalmazható empirikus módszereket. Tanulmányomban a térökonometria szemléletének bemutatására és néhány, a gyakorlatban gyakran használt térökonometriai modell és specifikációs teszt ismertetésére vállalkoztam. Az itt közöltek természetesen semmiképpen sem tekinthetők teljesnek. A térökonometria számos jelenlegi kutatási problémája (mint például a térbeli autokorreláció problémájának kezelése az egyéb, a lineáristól eltérő modellekben) kívül maradt az elemzés keretein. Befejezésként essék néhány szó a szoftver kínálatról.<sup>18</sup> A nemzetközi piacon létező legteljesebb programcsomag a térökonometriai és térstatisztikai elemzésekre kifejlesztett SpaceStat szoftver (*Anselin*; 1992, 1998). Gyakran használt még *Pace* és *Barry* (1998) Matlab rutinja a térökonometriai modellek maximum likelihood becslésére, valamint a Mathsoft S-Plus-hoz elkészített S+Spatialstats (*MathSoft*; 1996) a térbeli hiba autokorreláció modelljének becslésére.

#### IRODALOM

- ACS Z. – VARGA A. (2001): Térbeliség, endogén növekedés és innováció. *Tér és Társadalom*, 14. évf. 23–39. old.
- AOKI, M. (1994): New macroeconomic modelling approaches. Hierarchical dynamics and mean-field approximation. *Journal of Economic Dynamics and Control*, 18. évf. 865–877. old.
- AOKI M. (1996): *New approaches to macroeconomic modelling*. Cambridge University Press, Cambridge.
- ANSELIN, L. (1984): Specification tests on the structure of interaction in spatial econometric models. *Papers of the Regional Science Association*, 54. évf. 165–182. old.
- ANSELIN, L. (1988a): *Spatial econometrics: methods and models*. Kluwer Academic Publishers, Boston.
- ANSELIN, L. (1988b): Lagrange multiplier test diagnostics for spatial dependence and spatial heterogeneity. *Geographical Analysis*, 20. évf. 1–17. old.
- ANSELIN, L. (1992): *SpaceStat, a software program for the analysis of spatial data*. National Center for Geographic Information and Analysis, University of California, Santa Barbara, CA.
- ANSELIN, L. (1995): Local indicators of spatial association – LISA. *Geographical Analysis*, 27. évf. 93–115. old.
- ANSELIN, L. (1998): *SpaceStat Version 1.90*. <http://www.spacestat.com>
- ANSELIN, L. (2001): Spatial econometrics. In: *Baltagi, B. (szerk.): A companion to theoretical econometrics*. Basil Blackwell, Oxford, 310–330. old.
- ANSELIN, L. – FLORAX, R. (szerk.) (1995): *New directions in spatial econometrics*. Springer-Verlag, Berlin, 21–74. old.
- ANSELIN, L. – BERA, A. – FLORAX, R. – YOON, M. (1996): Simple diagnostics tests for spatial dependence. *Regional Science and Urban Economics*, 26. évf. 77–104. old.
- ANSELIN, L. – HUDAK, S. (1992): Spatial econometrics in practice. A review of software options. *Regional Science and Urban Economics*, 22. évf. 509–536. old.

<sup>18</sup> A szoftverkínálat részletes bemutatására lásd *Anselin* és *Hudak* (1992).



- ANSELIN, L. – REY, S. (1991): Properties of tests for spatial dependence in linear regression models. *Geographical Analysis*, 23. évf. 112–131. old.
- ANSELIN, L. – REY, S. (szerk.) (1997): A special issue on spatial econometrics. *International Regional Science Review*, 20. évf.
- ANSELIN, L. – VARGA A. – ACS Z. (1997): Local geographic spillovers between university research and high technology innovations. *Journal of Urban Economics*, 42. évf. 422–448. old.
- ATEN, B. (1996): Evidence of spatial autocorrelation in international prices. *Review of Income and Wealth*, 42. évf. 149–163. old.
- BALTAGI, B. – LI, D. (2001): Prediction in the panel data model with spatial correlation. In: *Anselin, L. – Florax, R. (szerk.): Advances in spatial econometrics.* (Kézirat.)
- BELL, K. – BOCKSTAEL, N. (1999): Applying the generalized moments estimation approach to spatial problems involving micro-level data. *The Review of Economics and Statistics*, 81. évf.
- BERON, K. – VIJVERBERG, W. (2001): Probit in a spatial context: a Monte Carlo approach. In: *Anselin, L. – Florax, R. (szerk.): Advances in spatial econometrics.* (Kézirat.)
- BOLDUC, D. – FORTIN, B. – GORDON, S. (1997): Multinomial probit estimation of spatially interdependent choices: an empirical comparison of two new techniques. *International Regional Science Review*, 20. évf. 77–101. old.
- BURRIDGE, P. (1980): On the Cliff-Ord test for spatial correlation. *Journal of the Royal Statistical Society B*, 42. évf. 107–108. old.
- CASE, A. (1991): Spatial patterns in household demand. *Econometrica*, 59. évf. 953–965. old.
- CASE, A. (1992): Neighborhood influence and technological change. *Regional Science and Urban Economics*, 22. évf. 491–508. old.
- CASE, A. – ROSEN, H. – HINES, J. (1993): Budget spillovers and fiscal policy interdependence: evidence from the States. *Journal of Public Economics*, 52. évf. 285–307. old.
- CASETTI, E. (1997): The expansion method, mathematical modelling and spatial econometrics. *International Regional Science Review*, 20. évf. 9–32. old.
- CLIFF, A. – ORD, J. (1981): *Spatial processes, models and applications.* Pion, London.
- CONLEY, T. (1996): *Econometric modelling of cross-sectional dependence.* PhD Dissertation. Department of Economics, University of Chicago, Chicago, IL.
- CRESSIE, N. (1993): *Statistics for spatial data.* Wiley, New York.
- DRISCOLL, J. – KRAAY, A. (1998): Consistent covariance matrix estimation with spatially dependent panel data. *The Review of Economics and Statistics*, 80. évf. 549–560. old.
- FUJITA, M. – KRUGMAN, P. – VENABLES, A. (1999): *The spatial economy.* MA, MIT Press, Cambridge.
- GRIFFITH, D. (1988): *Advanced spatial statistics.* Kluwer Academic Publishers, Boston.
- GETIS, A. – ORD, J. (1992): The analysis of spatial association by use of distance statistics. *Geographical Analysis*, 24. évf. 189–205. old.
- GREENE, W. (1993): *Econometric analysis.* MacMillan Publishing Company, New York.
- HAINING, R. (1990): *Spatial data analysis in the social and environmental sciences.* Cambridge University Press, Cambridge.
- HOLTZ-EAKIN, D. (1994): Public-sector capital and the productivity puzzle. *Review of Economics and Statistics*, 76. évf. 12–21. old.
- HUBERT, G. – CONSTANZO, C. – GALE, N. (1985): Measuring association between spatially defined variables: an alternative procedure. *Geographical Analysis*, 17. évf. 36–46. old.
- HEAGERTY, P. – LELE, S. (1998): A composite likelihood approach to binary spatial data. *Journal of the American Statistical Association*, 93. évf. 1099–1111. old.
- KELEJIAN, H. – ROBINSON, D. (1993): A suggested method of estimation for spatial interdependent models with autocorrelated errors, and an application to a county expenditure model. *Papers in Regional Science*, 72. évf. 297–312. old.
- KELEJIAN, H. – PRUCHA, I. (1998): A generalized spatial two stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *Journal of Real Estate Finance and Economics*, 17. évf. 99–121. old.
- KELEJIAN, H. – PRUCHA, I. (1999): A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review*.
- KÖRÖSI G. – MÁTYÁS L. – SZÉKELY I. (1990): *Gyakorlati ökonometria.* Közgazdasági és Jogi Könyvkiadó, Budapest.
- KRISTENSEN, G. (szerk.) (1996): *Symposium on the expansion method in the context of the family of models and methodologies with a focus on parametric variation, I-III.* Department of Economics, Odense University, Odense.
- KRUGMAN, P. (1991): Increasing returns and economic geography. *Journal of Political Economy*, 99. évf. 483–499. old.
- LAND, K. – DEANE, G. (1992): On the large-sample estimation of regression models with spatial-or network-effects terms: a two stage least squares approach. In: *Marsden, P. (szerk.) Sociological Methodology,* Jossey-Bass, San Francisco. 221–248. old.
- LESAGE, J. (1999): Spatial econometrics. In: *Web book of regional science.*  
<http://www.rri.wvu.edu/WebBook/LeSage/spatial/spatial.html>
- LUCAS, R. (1988): On the mechanics of economic development. *Journal of Monetary Economics* 22, évf. 3–42. old.
- LUNDAVAL, B. (szerk.) (1992): *National systems of innovation.* Pinter, London.
- MCMILLEN, D. (1992): Probit with spatial autocorrelation. *Journal of Regional Science* 32, évf. 335–348. old.
- MURDOCH, J. – RAHMATIAN, M. – THAYER, M. (1993): A spatially autoregressive median voter model of recreation expenditures. *Public Finance Quarterly*, 21. évf. 334–350. old.
- MURDOCH, J. – SANDLER, T. – SARGENT, K. (1997): A tale of two collectives: sulphur versus nitrogen oxides emission reduction in Europe. *Economica*, 64. évf. 281–301. old.
- MORAN, P. (1948): The interpretation of statistical maps. *Journal of the Royal Statistical Society B*, 10. évf. 243–251. old.
- MATHSOFT (1996): *S+SpatialStats user's manual for Windows and Unix.* WA: MathSoft, Inc., Seattle.
- NELSON, G. – HELLERSTEIN, D. (1997): Do roads cause deforestation? Using satellite images in econometric analysis of land use. *American Journal of Agricultural Economics*, 79. évf. 80–88. old.
- NELSON, R. (szerk.) (1993): *National innovation systems.* Oxford University Press, Oxford, New York.

- NEMES NAGY J. (1998): *Tér a társadalomkutatásban*. Ember–Település–Régió, Hilscher Rezső Szociálpolitikai Egyesület, Budapest.
- ORD, J. (1975): Estimation methods for models of spatial interaction. *Journal of the American Statistical Association*, 70. évf. 120-126. old.
- PAELINCK, J. – KLAASSEN, L. (1979): *Spatial econometrics*. Saxon House, Farnborough.
- PACE, R. – BARRY, R. (1998): *Spatial statistics toolbox 1.0*. Real Estate Research Institute, Louisiana State University, Baton Rouge, LA.
- PINKSE, J. – SLADE, M. (1998): Contracting in space: an application of spatial statistics to discrete-choice models. *Journal of Econometrics*, 85. évf. 125–154. old.
- PINTÉR J. (1991): A heteroszkedaszticitás diagnosztizálása. *Statistikai Szemle*, 69. évf. 1. sz. 16–36. old.
- ROMER, P. (1990): Endogenous technological change. *Journal of Political Economy*, 98. évf. S71–S102. old.
- TOBLER, W. (1979): Cellular geography. In: Gale, S. – Olsson, G. (szerk.): *Philosophy and geography*. Reidel, Dordrecht, 379-386. old.
- TOPA, G. (1996): Social interactions, local spillovers and unemployment. PhD Dissertation, Department of Economics, University of Chicago.
- UPTON, G. – FINGLETON, B. (1985): *Spatial data analysis by example*. Vol. 1: *Point pattern and quantitative data*. Wiley, New York.
- VARGA A. (1998): *University research and regional innovation: A spatial econometric analysis of academic technology transfers*. Kluwer Academic Publishers, Boston.

#### SUMMARY

Spatial econometrics is a subfield of econometrics dealing with the incorporation of spatial interaction and spatial heterogeneity into regression analysis. This paper provides a general introduction to spatial econometrics for the Hungarian professional audience by focusing primarily on the distinguishing characteristics of the methodology, on estimation of the most widely used linear spatial econometric models as well as on problems related to testing spatial autocorrelation.

### A HÁZTARTÁSI JÖVEDELMEK ÉS ADÓTERHEK RÉTEGMEGOSZLÁSA

DR. RÉVÉSZ TAMÁS

Elsősorban az 1998. évi Háztartási Költségvetési Felvétel és a legutóbb publikált, 1998. évi Ágazati Kapcsolatok Mérlege és háttértáblázatait, valamint a Nemzeti Számlák, a fogyasztásstatisztika, a Költségvetési Zárszámadás adatai felhasználásával a szerző 24 társadalmi-gazdasági réteg jövedelemit és kiadásait becsülte a makrostatisztikai adatokkal összhangban, a munkajövedelmeket és fogyasztásokat részletes szakágazati bontásban, a többi jövedelmeket és elvonásokat pedig a megfelelő jogcímek szerinti bontásban. Erről a soklépéses, számos transzformációt, korrekciót és módszertani megfontolást igénylő eljárásról a cikk részletesen beszámol, az eredmények mellett az adathiányokból, inkonzisztenciákból, módszertani eltérésekből eredő problémákat, és azoknak a szerző által történt kezelését is megvilágítva.

Az eredményül kapott adatbázis legfontosabb részeit, valamint az adatbázis alapján számított különféle adóteher- és jövedelemeloszlási mutatókat, egy főre számított jövedelmeket és adókat részletes táblák mutatják be. Ezek az itteni eredményeknek korábbi és külföldi vizsgálatok eredményeivel való összehasonlítását is tartalmazzák. A szerző elemzést, következtetési lehetőséget is felvázol, legfontosabb megállapításai közül említhetjük az adóprogresszió gyengeségét, a gyermekes háztartások túlzott adóterhelését, a középosztály alacsony jövedelemrészesedését, a jövedelemegyenlőtlenség növekedését, az egyes rétegek munkakötődését különféle ágazatokhoz.

A cikk befejező részében felvázolja a továbbfejlesztés lehetséges fő irányait.

TÁRGYSZÓ: Jövedelem. Háztartás-statisztika. Adózás. Jövedelemeloszlás.

A piacgazdasági átalakulás során lezajlott nagyarányú társadalmi-gazdasági szerkezetváltozások és az aggregált nemzeti jövedelem utóbbi években megfigyelt erőteljes növekedése az egyes társadalmi rétegeket meglehetősen eltérően érintették. A rétegek helyzetéről csak részleges ismereteink vannak, noha ez a gazdaságpolitika szempontjából is rendkívül fontos. Tekintettel a magyarországi magas adókulcsokra érdemes egy olyan átfogóbb képet felvázolni, amely a rétegek helyzetét alapvetően befolyásoló adóterhek megoszlását helyezi előtérbe. Az adóterhelés elemzését célszerűnek látszik egy makrogazdasági helyzetképbe illesztve végezni. A jelen tanulmányban ilyen széles kitékintésre azonban nem vállalkoztam részben terjedelmi okok, részben pedig amiatt, mert más kormányzati dokumentumokban és publikációkban a makrogazdasági helyzetkép és összefüggések jelentős része megtalálható.

A jövedelmek és az adók statisztikai elemzése ideális esetben egy, a gazdaság szereplőinek bevételeit és kiadásait teljeskörűen, és egymással konzisztens módon ábrázoló, ún. Társadalmi Elszámolási Mátrix (angol rövidítéssel: SAM) összeállításából és elemzéséből indul ki. Ezt azonban a magyar statisztikai rendszer még nem tudja összeállítani. Kutatóként azonban különféle hipotézisekkel már több ízben is vállalkoztam ilyen mátrixok összeállítására. Az 1990. évre és 1998-ra vonatkozó, a háztartási szektort 10-10 rétegre bontva ábrázoló SAM-okat az olvasó megtalálhatja a *Révész–Zalai* (1993) és *Révész* (2001) munkákban.

A jövedelem- és adóteher-megoszlásnak a piacgazdaság újraéledésekor állapotát érdemes összevetni a piacgazdaság mintegy 60 évvel azelőtt (a győri fegyverkezési programmal) kezdődött fokozatos felszámolása előtti állapotával. Ezt az összehasonlítást teszi lehetővé az akkori adóteher-megoszlásról *Matolcsy Mátyásnak* 1938-ban megjelent úttörő jellegű tanulmánya (*Matolcsy*; 1938).

A történelmi léptékű összevetésen túlmenően természetesen érdemes az adóteher-megoszlásra kapott eredményeinket más országok adózási helyzetével, és az adóteher-megoszlásra vonatkozó más vizsgálatok eredményeivel is összehasonlítani. A tanulmányban erre csak korlátozottan vállalkoztam, inkább csak a lehetőségeket és az összehasonlításhoz szükséges módszertani megfontolásokat igyekeztem felvázolni.

Az egyes rétegek jövedelmei és az adóteher megoszlásának számításánál alapvetően az 1998. évi Háztartási Költségvetési Felvétel (HKF), az 1998. évi szakágazati bontású Ágazati Kapcsolati Mérlegek (ÁKM), a Nemzeti Számlák (NSZ), valamint a költségvetési Zárszámadás kötetek adataira támaszkodtam. Elvileg lehetne úgynevezett *modell-háztartásokra* (különböző feltételezett demográfiai szerkezetű és jövedelmi-adózási helyzetű háztartásokra) is kiszámítani az adóterhelést és annak változását, de mivel ezzel korábbi publikációimban már foglalkoztam, (lásd például *Newbery–Révész*; 2000) most *Matolcsy* (1938) empirikus módszerét követve csak a ténylegesen megfigyelt adókkal és jövedelmekkel foglalkozom.

A jövedelem- és adómegoszlást többféle rétegbontásban is kiszámítottam. A számszerű eredményeket megkísérlem értékelni, és néhány gazdaságpolitikai következtetési lehetőségre rámutatni. Végül összefoglalom a továbblépés, az esetleg célszerű korrekciók és finomítások, valamint a számítások kiterjesztésének főbb irányait és teendőit.

#### A HÁZTARTÁS-STATISZTIKAI ADATOK FELHASZNÁLÁSA AZ ELEMZÉSBEN

Az alternatív rétegbontások lehetőségének megteremtése végett a háztartás-statisztikáról rendelkezésre álló publikáción (*Családi költségvetés*; 1999b) túlmenően az 1998. évi HKF számunkra fontos egyedi adatait az ECOSTAT Gazdaságelemző és Informatikai Intézet segítségével SAS-adatállományként is felhasználtam. Ezeknek az adatoknak egy erre a célra kifejlesztett SAS-programmal való feldolgozása révén lehetővé vált a HKF számunkra megfelelő tartalmú, rugalmasan módosítható kategóriáinak képzése és számszerűsítése.<sup>1</sup>

<sup>1</sup> A HKF SAS-programmal való feldolgozásában való segítségért *Nagyné Pozsgai Máriának* és *Takács Tibornak* tartozom köszönettel.

A HKF-adatoknak a makrostatisztikai (perem-)adatokhoz való körültekintő illesztésével jött létre az elemzés átfogó, konzisztens, részletes és rugalmasan módosítható adatbázisa. A folyamat teljes leírása, és a háttéradatok részletes bemutatása természetesen vastkos kötetet tölthetne meg, ezért a továbbiakban csak a legfontosabb lépések és tanulságok rövid összefoglalására vállalkozhatunk.

A későbbi elemzési lehetőségek szélesítése végett a megkérdezett 10 144 háztartás HKF-beli adatait (természetesen a szokásos demográfiai–területi mintavételi valószínűségek reciprokával való felszorzás után) – négy ismérv (aktivitás, településtípus, gyermekszám, jövedelmi helyzet) lehetséges állapotainak Descartes szorzata révén kapott – 24 rétegre összesítettem. E rétegek némelyike önmagában meglehetősen kicsi, de elkülönített kimutatásukra szükség volt, mivel más-más aggregációknál máshova kellett őket sorolni. A túl részletes bontásból származó számadatok az egy csoportot képviselő átlagosan kevés megfigyelési egység miatt nem megbízhatók. A dezaggregálásra sokszor nem a részletek iránti kíváncsiság miatt van szükség, hanem azért, hogy a számítások ellenőrizhetőbbek, az aggregált eredmények megbízhatóbbak legyenek.

A jövedelmi helyzet esetében konkrétan az egy főre jutó nettó jövedelem nagysága szerinti három csoportot képeztem. A három jövedelmi réteg elhatárolását – figyelembe véve a magas jövedelműek alulreprezentáltságát és a jövedelmek egy részének meg nem jelenését (elhallgatását, kifejejtését stb.) az alapadatokban – az egy főre jutó jövedelmek mintabeli eloszlásának 40 és 80 százalékos határán végeztem el. Így a két alacsonyabb jövedelmi csoport együtt a népesség 80 százalékát teszi ki, ami jól összehasonlítható Matolcsy hivatkozott alapművének I. jövedelmi csoportjával, a „szerénysorúak”-kal, akik a népesség 81,2 százalékát alkották. A magasnak tűnő 80 százalékos határ kijelölésének pedig az is a szempontja volt, hogy hazánk európai mércével alacsony jövedelmi színvonal mellett „jómódú”-nak csak a társadalom kisebb része tekinthető, a természetesnek tűnő egy-egyharmadnyi szegény–közepes–jómódú felosztással szemben.

#### *A jövedelmek becslése*

Mindenekelőtt tisztázni kellett, hogy a HKF-ben milyen jövedelemfajtákra van adat, illetve ezek közül melyeket célszerű figyelembe venni. A célszerűséget elsősorban a más elemzésekkel való összehasonlíthatóság, az adatok és becslési eljárások megbízhatósága, valamint adóelméleti megfontolások alapján lehet meghatározni.<sup>2</sup> Mindezek alapján úgy döntöttem, hogy az NSZ-ben a keresetekben, valamint a házilagosság építkezés értékén kívül az ún. „vegyes” (vállalkozási) jövedelemben megjelenő részek figyelembe vehetők az adóterhelési mérték számításánál, az NSZ-ben rendre 492, 54, illetve 1288 milliárd forint értékűnek becsült jövedelmet hozó imputált lakásszolgáltatás, házilagosság építkezés és (a főleg a személyi jövedelemadót nem fizető gyermekeknek és nyugdíjasoknak nyújtott) természetbeni társadalmi juttatások (oktatás, egészségügyi ellátás stb.), valamint az NSZ-ben figyelembe nem vett háztartási munka és egyéb (adóalapként egyébként sem szóba jöhető, esetleg nem is törvényes eredetű) tételek viszont nem.

<sup>2</sup> Például amennyiben a természetbeni jövedelmek az adott termékcsoportra költendő saját kiadásokat mérséklék, amnyiban az így megtakarított összeggel elvben növelhető az adóterhelésnél számításba vehető adóalap. E helyettesíthetőségre vonatkozóan azonban nem áll rendelkezésre adat.

Az így definiált jövedelmek jelentős részét képező kereseteket szakágazati eredet szerint is megbontottam. Erre a felbontás nyilvánvaló foglalkoztatáspolitikai fontossága mellett azért volt szükség, hogy egyes ágazatokból származó kereseteknek a HKF-beli jellegzetes alulreprezentáltságát megfelelő módon kezelhessük.

Konkréten, első lépésben a keresetek *szakágazati eredet*<sup>3</sup> szerinti felosztását a háztartások egyes személyeinek szakágazati hovatartozása alapján végeztem el.<sup>4</sup> Ez természetesen a mellékkeresetek tekintetében esetenként téves besorolást eredményez, ezek a tévedések azonban nem túl jelentősek, és feltehetően a rétegek szintjén jórészt kiegyenlítődnek.

Az elemzés következő lépésében tehát már csak e torzítás kiküszöbölésére is az egyes rétegeknek a háztartás-statisztikából becsült szakágazati kereseteit olyan arányban korrigáltam, hogy a rétegeknek az adott szakágazatból származó összes munkajövedelme megegyezzen a nemzeti számlabeli értékkel.

A közvetlenül a mezőgazdaságba sorolt háztáji termelésen kívüli vállalkozási jövedelmeket hasonló módon osztottam szét szakágazatokra. Itt átlagosan nagyobb arányú makrostatisztikai kiigazításra volt szükség, és a relatív hiba is nagyobb lehet, a vállalkozási jövedelmek viszonylag alacsony aránya miatt azonban ez az összképet aligha módosítja érdemlegesen.

A háztartások pénzbeli társadalmi juttatásokból származó jövedelmei kevésbé voltak problematikusak. Ez részben az érintetteknek a mintában való nagyobb arányú részvételéből, és bevallási hajlandóságából ered. Így a HKF-ben a legnagyobb tételek, a nyugdíj, a családi pótlék, a gyés és a munkanélküli jövedelempótló támogatás majdnem teljesen a makrostatisztikai adatokkal egyezően jelenik meg.<sup>5</sup> Legkisebb mértékben, egynegyedes, illetve egyhatedes arányban a táppénz és a szociális segélyek jelennek meg a HKF-ben, külön oknyomozás nélkül is sejthető okokból. Mindenesetre e tételeket is arányosan hozzáigazítottam az egyébként a nemzeti számla kiadványból, a Zárszámadásból, illetve a Szociális statisztikai évkönyvből vett makrostatisztikai adatokhoz mint előírt összesenekhez.

Természetesen az arányos kiigazításoktól eltérően, lehetne a hiányzó összegeket az egyes rétegeknél eltérő arányúnak feltételezve, rétegenként eltérő szorzóval kiigazítani. Egy ilyen kísérlet található *Szabóné* (1996) cikkében. E cikk szerzője azonban a korrekciót az egyes rétegek összjövedelmén végezte el (mégpedig a legalsó decilisnél 25 százalék, majd egyenletesen emelkedő mértékben a legfelső decilisnél pedig már 45 százalék hiányzó jövedelmet feltételezve). Efféle korrekcióra azonban nem vállalkoztam, részben információhiány, részben pedig amiatt, hogy az egyes jövedelemelemek külön-külön való kiigazítása részben hasonló eredményre vezet, legalábbis amennyiben a nagyobb mértékben hiányzó jövedelmek a magasabb jövedelműeknél koncentráltabban jelentkeznek (főleg mivel ők az átlagnál nagyobb mértékben a HKF-ben kevésbé reprezentált szolgáltatásokban tevékenykednek). Egyébként az említett szerzőnek a jövedelmek nagysága szerint folyamatosan emelkedő kipótlási módszerével ellentétben a

<sup>3</sup> A NSZ-ban található 57 ágazatból 4 alágazatot, az integrált szénbányákat, a MOL kitermelési tevékenységét, az alapanyag-vegypart és a gázelosztást kivonva összesen 61 ágazatot szerepeltettem az elemzésben.

<sup>4</sup> A személyek ágazati besorolásának egyes esetekben a HKF-ben is (feltehetően a vállalat szervezeti hovatartozása helyett a tevékenység jellege alapján történt) téves voltát az elsősorban érintett szakágazatokon belüli belső átcsoportosításokkal részlegesen korrigáltam.

<sup>5</sup> Bár az utóbbi tételbe, esetleg tévedésből, a csak 50 százalékosan reprezentált munkanélküli járadék egy részét is beírták.

*Világbank* (1996) magyar országtanulmánya alapján „az elképzelések szerint” a hiányzó jövedelmek aránya a jövedelmek U-alakú függvénye, azaz nem egyenletes. Ez is alátámasztja azt a megoldásomat, ami a kiigazítást nem aggregált módon, hanem összetevőnként külön-külön valósította meg.

A jövedelmek között makrostatisztikai értékhez való *kiigazítás nélkül* még hozzászámítottam a HKF-ben „felvett kamat, osztalék” és „egyéb jövedelem” néven megjelenő összegeket. Az egyéb jövedelmek<sup>6</sup> meglehetősen heterogén kategóriája esetében a makrostatisztikai referenciaértékek hiánya miatt nem lehet tudni, hogy a HKF-ben megjelenő 40 milliárd forint körüli összeg a valóságosnak mekkora hányada.

A kamatokra és osztalékokra ugyan van a nemzeti számlákban referenciaadat, ezek azonban csak a nominális hozamok, reáljövedelmet a magas infláció, illetve a tőzsdei árfolyamok esése miatt csak kis részben tartalmaz. Kompromisszumos megoldásként a felvett részt vettem figyelembe, ami például abban az esetben is helyes megoldás, ha a rétegek éppen csak a reálhozamot veszik fel, illetve írták be a kérdőívre.

Végeredményben az eddig tárgyalt jövedelmek összegét tekintve az átlagosan 1,79-szoros arányos kiigazítások a 24 réteg közül a 12 inaktív rétegnek a bruttó jövedelmét 1,15–1,6 közötti arányban növelte, a 12 aktív rétegnek pedig 1,58–2,43 közötti arányban. A legnagyobb arányú kiigazítás nem meglepő módon éppen az (684 ezer fős) aktív városi jómódú gyermektelen háztartásoknál történt. Összességében, az alacsony jövedelműek jövedelme átlagosan 1,63-szorosra, a középosztályé 1,69-szorosra, a jómódúaké pedig 2,02-szorosra nőtt a kiigazítástól várt progresszív módon.

A háztartások közötti transzferek is szerepelnek a HKF-ben, és ezek elenyészőnek mutatózó egyenlegét mint a nettó jövedelem egyik komponensét a háztartás-statisztikán alapuló más elemzésekkel való összehasonlíthatóság végett fel is használtam a jövedelmi csoportok képzésénél. Az alacsony értékek, az adatok megbízhatósági problémái, valamint elméleti megfontolások miatt a háztartások közötti transzfereket nem vesszük figyelembe a jövedelemszint, az adóterhelés és ezek megoszlásának elemzésénél.

#### *A jövedelmekhez kapcsolódó elvonások becslése*

Az előző részben számított, valós (az NSZ-szel konzisztens) mértékű jövedelmeket természetesen csak a szintén valós mértékű adóterheléssel lehet értelmesen összehasonlítani. Ezért a jövedelmeknél használt arányos kiigazítást alkalmaztam a HKF-ben megjelenő elvonási tételeknél is. Ezek közül a nyugdíjjárulék és a gépjárműadó alig szorult korrekcióra, míg a személyi jövedelemadónak csak a kétharmada, az NSZ-ben a háztartásoknál és transzferként elszámolt (tehát például idegenforgalmi adó és vállalati illeték nélküli) egyéb illetékeknek az egynegyede, helyi adókból pedig az egyhetede jelent csak meg a HKF-ben. Mivel azonban összességében sem nagyon jelentős e két tétel, egyelőre nem foglalkoztam azzal, hogy a nagyobb arányú felszorzás okozta esetleges torzítás hogyan csökkenthető. Az egyéb illetékeket a kiadásoknál tárgyaljuk.

Az egyenes adókat ténylegesen megfizetett összegükben vesszük számításba, így az ún. „effektív” adóterhelés számítható. Az adópolitika szempontjából később célszerű len

<sup>6</sup> Tételeken a HKF-ben szereplő, a háztáji és illetményföld helyett kapott pénz, az életjáradék, a kárpótlási jegyért kapott életjáradék, a bérbeadásból származó bevétel, az egyéb külföldi jövedelem, a „más jogcímen” kapott jövedelem, és a szűkebb értelemben vett egyéb jövedelem összegét neveztem egyéb jövedelemnek

ne a bruttó adókat és az adókedvezményeket<sup>7</sup> külön kimutatni. Erre egyelőre az elméleti nehézségek<sup>8</sup> és amiatt nem vállalkoztam, mert a HKF-ben nincs jogcímenként bontott információ az adókedvezményekre.

#### *A kiadásokhoz kapcsolódó adók becslése*

E fejezetben a különféle, az NSZ-ben termékadóként elszámolt tételekből az egyes rétegeket terhelő összegek meghatározásának módszerét tekintjük át.

Természetesen az NSZ (és az azt követő ÁKM) szerinti termékadó-fogalom is keltethet bizonyos hiányérzeteket. Például a Zárszámadásban megjelenő illetékeket részben fogyasztásnak, részben transzfernek, részben pedig tőkeadónak, termékadónak csak a beruházásra jutó részt tekintik. A legnagyobb illetékfajtát, a repülőtéri illetéket azonban már a Zárszámadás sem illetékként, hanem szolgáltatási („működési”) díjbevételeként mutatja ki.

A háztartásokat terhelő illetékek makrostatisztikai értékének a KSH egy munkatáblája alapján a 20 milliárd forintos lakossági beruházási illeték, a transzferként elszámolt 3,8 milliárd forint, valamint a tőkeadóként feltüntetett 4,3 milliárd forint összegét tekintettem (azaz a szolgáltatási díjként elszámolt 5,6 milliárd forint kivételével valamennyi lakosságra vonatkozó illetéket). Mivel azonban a HKF külön kiadási tételként szerepelteti a járműadókat és -illetékeket, a gépjármű-átírási illeték (frissebb adat híján az 1995. évi Zárszámadásban közölt) 6,6 milliárd forintos összegét ez utóbbi kategória makrostatisztikai referenciaértékébe csoportosítottam át. A járműhöz kapcsolódó és nem kapcsolódó adók és illetékek makrostatisztikai összegeit az adott kategóriára vonatkozó, HKF-ben szereplő rétegadatok arányában osztottam szét rétegekre.

A háztartásokat terhelő adók részben a lakossági beruházások révén folynak be. A beruházási illetékek kezelését már említettük (azaz, hogy tekintve a kisebb lakások illetékmentességét, az összes illetékekkel, és nem valamely lakásberuházási tétellel arányosan osztottam szét rétegekre). Ennél azonban nagyobb tétel a Zárszámadásban 25 milliárd forintos („előzetes”) összegű (nettó) lakásberuházási áfa (Láfa). Ez jól egyezik az ÁKM állóeszköz-felhalmozási termékadó, az NSZ állóeszköz-felhalmozási, a Lakásstatisztikai évkönyv (Láfa nélküli) magánlakás-beruházási és a Zárszámadás önkormányzati nettó felhalmozási áfa adataiból – az összes adó és az egyéb tételeket terhelő adók különbségeként – kétféleképpen is meghatározható maradványértékekkel.

Ezt a 25 milliárd forintot a lakásberuházási kiadások arányában osztottam szét rétegekre. Sajnos a HKF-ben a lakásberuházásoknak csak mintegy egyharmada jelenik meg (az építkezők kevésbé érnek rá kérdőíveket tölteni, az építkezés, illetve költözés időszakában pedig gyakran nem is lehet megtalálni őket), és elképzelhető, hogy a hiányzó kétharmad nem ugyanolyan arányban oszlik meg a szóban forgó rétegeink között, mint a megjelenő egyharmad rész. Tekintve azonban a 25 milliárd forintos összeg nem túl nagy súlyát, egyelőre a lehetséges korrekciókkal nem foglalkoztam.

<sup>7</sup> Az APEH honlapján olvasható tájékoztató szerint a személyi jövedelemadó-kedvezmények 1998-ban 282 milliárd forintra rúgtak, amiből 154 az alkalmazotti adójóváírás, 40 a nyugdíjszerű jövedelmek adókedvezménye, 36 a nyugdíjjáruulék utáni, 21 a kiegészítő nyugdíjpénztári, 19 milliárd forint a befektetési és biztosítási rész. A HKF-ből 153 milliárd forint adókedvezmény számítható.

<sup>8</sup> Például, hogy különösen a vállalkozási tevékenységre tekintettel nehéz megmondani, hogy mennyi lenne a normális (nem kedvezményes) adóalap és adókulcs.



Az egyes fogyasztási kiadási tételekhez kapcsolódó adók közvetlenül nem jelennek meg a HKF-ben. Ezek valós, makrostatisztikai összegeinek rétegek közötti elosztásához az adót viselő termékek HKF-beli fogyasztási összegek rétegek közötti megoszlási arányait használtam. Természetesen itt is lehet arról szó, hogy az egyes javak rejtett (a HKF-ben meg nem jelenő) fogyasztása az egyes rétegeknél eltérő arányú, a jelenlegi kiindulási alapnak szánt számításaimban azonban az esetleges eltérésekkel nem foglalkoztam, mivel itt is jó esély van arra, hogy a sokféle cikksoport összesítésénél az egyes cikksoportoknál mutatózó, eltérő irányú korrekciók jelentős részben kiegyenlítődnek.

Az eljárás során először a fogyasztási kiadások szakágazati eredet szerinti bontásban való transzformációjára volt szükség, mivel legteljesebben ebben a bontásban lehetett részletesen meghatározni az egyes termékek tényleges (effektív) adókulcsát, az 1998. évi nagy ÁKM és háttértáblázatai fogyasztási és arra jutó termékadók vektoraiból.<sup>9</sup>

A HKF fogyasztási kategóriáinak rétegek szerint bontott, az össznépeségre felszorozott adatait az 1998. évre a KSH-ból származó fogyasztásstatisztikai adatok, az ÁKM CD-ROM változatához mellékeltem ún. levonómátrixok között található kereskedelmi és szállítási árés adatok, valamint egy korábban kidolgozott fogyasztástranszformációs mátrix alapján osztottam szét szakágazatokra (beleértve az árés leválasztását és a kereskedelembe, illetve a szállítás ágazatba való átrakását is).

Természetesen meg kellett határozni, hogy az ÁKM-ben a hazai fogyasztáson elszámolt adókból mekkora a rezidens (hazai) háztartásokra jutó rész, és mekkora a beutazó turisták fogyasztását terhelő összeg. Sajnos, mint ismeretes, a beutazó turisták fogyasztására nincsenek megbízható statisztikai adatok még az összkiadásuk tekintetében sem, de különösen termékszerkezetét illetően (lásd *Hüttl*; 2000). A problémák ellenére a beutazó turisták összkiadása tekintetében az MNB új fizetésimérleg-számát fogadtam el, azzal a hipotézissel, hogy az NSZ-ben figyelembe vett értékhez képesti (217 milliárd forint) többlet a magánfogyasztásra jut. Az így becsült mintegy 730 milliárd forintos turistaexportot a KSH 1994. évi (*Külföldiek...*; 1996), valamint a Gfk Hungária 1998-as felmérése (*Gfk*; 1999) adatai, *Horváth* (1999) tanulmánya, valamint a WTTC „Turizmus Szatellit Számola” c. tanulmánya<sup>10</sup> alapján osztottam szét ágazatokra.

Az egyes termékekre jutó adókból a turistákra jutó részt a fogyasztásukkal arányosnak tekintve, majd a kapott értékeket összesítve 117 milliárd forint értékű adódott, ami a turistaköltség 16 százaléka.<sup>11</sup>

A szeszes italokat terhelő, a Zárszámadás szerint 41 milliárd forintos fogyasztási adó (+25 % áfa) beutazó turistákra, illetve hazai rétegekre jutó összegeinek meghatározásához azonban az ÁKM kétszámjegyű szakágazati bontása (konkrétan az élelmiszerek és italok gyártása szakágazat adózás szempontjából heterogén jellege miatt) nem volt megfelelő. A realitásabb rétegek közötti elosztása végett itt kivételesen az adókat még a szakágazati besorolás előtti, a HKF-ben még élvezeti cikkeként szereplő fogyasztások arányában osztottam szét rétegekre és a beutazó turistákra. Ez utóbbiak részesedését pontosabb információ híján az élelmiszer-fogyasztásuk arányával azonosnak tekintettem.

<sup>9</sup> Továbbá hasonló módon lehetett a terméktámogatásokat is meghatározni.

<sup>10</sup> A WEFA/WTTC Inc. által a Magyar Turizmus R.T. részére „Simulated Tourism Satellite Account for Hungary” címmel 2000 júliusában benyújtott, egyelőre nem publikus tanulmány.

<sup>11</sup> Ez összehasonlítva a *Horváth* hivatkozott tanulmányában 17 százalékosnak becsült közvetlen vám- és adótartalommal, meglepően jó egyezést mutat.

Hasonló, de ellentétes előjelű probléma vetődött fel a gyógyszerfogyasztások esetében is. Ezek adót nem tartalmaznak, tehát a 24-es számjelű, „vegyi anyag, termék gyártása” elnevezésű szakágazatra jutó adókat a (arányos kiigazítással korrigált HKF-beli) gyógyszerfogyasztási adatok leválasztása után maradó fogyasztási összegek arányában osztottam szét rétegekre.

Általában a szakágazati bontás a cikkszoportos bontáshoz képest akkor hátrányosabb, ha ugyanazon szakágazatba eltérő adótartalmú, ráadásul az egyes rétegek által eltérő arányban fogyasztott termékek kerülnek. Kisebb mértékben ez áll fenn még a kőolaj-feldolgozás, a posta-távközlés, a szárazföldi szállítás, a szórakoztatás–kultúra–sport szakágazatok esetében is. Ezeknél a kapott eredményeket érdemes felülvizsgálni az eredeti, cikkszoportos adatok alapján (bár gyakran még ezek felbontása sem eléggé teszi lehetővé az eltérő adótartalmú termékek különválasztását).

A vámok elemzésbe való bevonásával szemben azonban bizonyos elvi ellenvetések tehetők. Amennyiben ugyanis a vámok protekcionista jellegűek, úgy a fogyasztó a hazai termék világpiacínál magasabb áron való megvásárlásával is – a vámmal azonos mértékű – burkolt adót fizet az egyébként versenyképtelen hazai termelő fenntartása végett. Ez tehát elvezet ahhoz az általános elméleti állásponhoz – aminek számszerűsítési módszeréről lásd például *Adelman et al.* (1991) cikkét –, ami minden, a piaci egyensúlytól való állami áreltérítést elvonásnak, illetve támogatásnak tekint. Bár ezen áreltérítő hatásokra vonatkozóan korábban magam is beható elemzéseket végeztem (*Révész–Zalai*; 1992), e tanulmányban erre nem vállalkozom. Ezért a jelen tanulmányban a fogyasztásra eső, a magánimporton beszedett 1 milliárd forintos értéket (lásd *Zárszámadás* 113. old.) és az ÁKM-ben a magánfogyasztáson elszámolt 41 milliárd forint vámot osztottam szét rétegekre. A magánimport csekély összegű vámját egyszerűen a külföldi üdülési kiadások arányában osztottam szét rétegekre. Bár e kiadások (feltehetőleg) nem tartalmazzák az elvámolt cikkeket, valószínűleg – legalábbis a rétegösszesenek szintjén – erősen korrelálnak azzal.

Ezután a cégimport 41 milliárd forintnyi vámját osztottam szét először ágazatokra, majd fogyasztásarányosan rétegekre. Noha a fogyasztásra jutó vámok ágazati bontása nem állt rendelkezésre, az összes felhasználásra jutó vámoké azonban igen. Ebből, és az ÁKM-ben felhasználónként bontott vámokból, és a vámszabályokból viszonylag megbízhatóan lehet becsülni az egyes importtermékek fogyasztására jutó vámokat. Konkrétan egy biproporcionális kiigazítást jelentő RAS-feladatot oldottam meg (e módszerről lásd például *Lecomber* (1995) kritikai áttekintését), ahol a felhasználásokon belül a termelőfelhasználást (további bontás nélkül), a magánfogyasztást, és az állóeszköz-felhalmozást különböztettem meg, majd az ezekre elszámolt vámokhoz és az ágazati bontásban megadott vámokhoz mint peremekhez (előírt sor- és oszlopösszesenek) igazítottam a vámok felhasználás és ágazat szerint bontott (az importtal képviselt) indulómátrixát.

Az így kapott, a magánfogyasztásra jutó, ágazatonként bontott vámokat az egyes ágazatok magánfogyasztásának arányában osztottam szét rétegekre és a külföldi turistákra. Az importfogyasztásokkal arányos elosztásra nem volt mód, mert a rétegek importfogyasztásaira vonatkozóan nem volt adat. A fogyasztásarányos elosztás tehát impliciten feltételezi, hogy az egyes rétegek egyazon terméksoporton belül azonos arányban vásároltak importcikkeket. Ez ugyan – különösen a turisták esetében – egyes terméksoportokban aligha áll fenn (a külföldiek néha szuvenírként kifejezetten magyar terméket vásá-

rolnak, máskor viszont ódzkodnak a számukra ismeretlen hazai termékektől), az esetleges hibák az összesítéskor valószínűsíthetően kiegyenlítődnek. Ezzel a módszerrel egyébként a turisták vásárlásaira jutó vám mintegy 2,7 milliárd forintra tehető. Ez a várakozásoknak megfelelően fogyasztási arányuknál alacsonyabb arányt jelent.

A rezidens háztartásokra jutó termékadóknak az ismertetett módon ágazatonként meghatározott összegei rétegek közötti elosztásánál elvileg a rétegek fogyasztásából a külföldön történt kiadásokat le kellett volna vonnunk. Erre azonban többféle okból nem volt szükség, illetve lehetőség. Egyrészt, a külföldre utazó háztartások érhetően kisebb arányban vannak a háztartási mintában képviselve, másrészt a kinti költségeket csak egy összegben, mint „külföldi üdülés” kiadásait írták be. A tartós fogyasztási cikkek külföldön történt vásárlása természetesen az adott termékre beírt kiadásban is jelentkezhet, de ez 1998-ban már nem lehetett túl jelentős, illetve a rétegek közötti fogyasztási arányokat nem módosíthatta lényegesen. Közismert például az import használtautók – ráadásul általában cégautóként való – behozatalakor jelentett alacsony érték, azaz vám- és adóalap. Mint láttuk, az összes magánimporton a Zárszámadás szerint mindössze 1 milliárd forint vámot szedtek be.

A háztartások pénzügyi megtakarítását, illetve az abból származó jövedelmeit a kamatadó 1994. évi eltörlése óta csak az osztalékadó terheli. Az egyes rétegek pénzügyi megtakarításaira sem a háztartás-statisztikából, sem más adatforrásból nem rendelkezünk adatokkal. A megtakarításokat leginkább az infláció mint burkolt adó sújtja, ezt azonban a magasabb nominális kamatok rendes körülmények között ellensúlyozzák. A készpénzre természetesen ez nem igaz, így a háztartások készpénzállományából erre az inflációs adóra (vagy nemzetközi szakkifejezéssel „seigniorage”-ra) elvileg lehetne következtetni. Sajnos azonban ezt a háztartási naplóban még szereplő adatot a SAS-típusú adatállomány már nem tartalmazza. Igaz, a kis mintaelemszám miatt már problémát jelenthet, hogy az egyes háztartások más-más hónapban vezettek naplót, jóllehet az egyes hónapokban a lakosság készpénzállománya meglehetősen eltérő szokott lenni. Ezért jelen tanulmányban nem vettem figyelembe a pénzügyi megtakarításokat terhelő explicit vagy burkolt közvetlen adókat.

A szakirodalomban azonban a megtakarítások közvetett adótartalmát is kiszámítják néha (lásd például *Newbery–Révész*; 2000), arra hivatkozva, hogy a megtakarítás az elhalasztott fogyasztás, és előbb vagy utóbb fogyasztási adó terheli. Mivel azonban a pénzügyi megtakarítások csak részlegesen és nem megbízható (a jövedelmekkel és fogyasztásokkal az évesített adatok szintjén nem is konzisztens) módon szerepelnek a HKF-ben, konkrét rétegelosztást nem tudunk számítani.

#### *Záró megjegyzések az adatbecsléshez*

Érdeemes a vázolt módszert összevetni más hasonló kísérletekkel (*Gál et al.*; 2000). Az ÁKM-alapú módszerem (amit természetesen az ÁKM 2001. februári megjelenése előtt nem lehetett alkalmazni) előnye – azokon az említett tényeken kívül, hogy a fogyasztásra jutó részt mutatja (ellentétben a Zárszámadásban szereplő, felhasználási területenként nem bontott adatokkal) és ágazati bontásban –, hogy minden termékadót magába foglal, beleértve a vámokat, az iparüzési adót, az idegenforgalmi adókat, a szeszadót, a kulturális járulékot és játékadót, valamint a környezetvédelmi termékdíjat is. Bár az egyes

adófajtákat elkülönítve nem mutatja be, az adómértékekből és az ágazati szerkezetből az eszerinti bontás is viszonylag könnyen előállítható. Erre azonban a jelen tanulmányban nem volt szükség (az áfa-visszaigénylésre úgysem jogosult fogyasztónak mindegy, hogy az adott összeget fogyasztási adó, vagy forgalmi adó címen szedik be tőle). Az ÁKM-kompatibilis adatok további nagy előnye, hogy az ÁKM segítségével a közvetlen adótar- talmon túlmenően a közvetett, a szóban forgó termékek előállításánál felmerülő adó- vonzatokat is számíthatjuk.

Ezzel kapcsolatban egyes szerzők megkísérik a vállalati adókat is a háztartásokra ter- helni, nincs egyértelmű álláspont azonban arról, hogy ez esetben a vállalati adók milyen mértékben terhelik a vállalat tulajdonosait (a „tőkét”), munkavállalóit, illetve fogyasztóit, esetleg beszállítóit (itt is érvényes az áreltérítésekkel kapcsolatban vázolt elméleti állás- pont). Mindenesetre célszerű részleteiben és alternatív hipotézisekkel vizsgálni a kérdést, de erre a nagyvolumenű munkára egyelőre nem vállalkoztam.

Természetesen még számos más, elméletileg is nehéz probléma vetődött fel. Ezek kö- zül csak jelzésértékűen az alábbiakat emelem ki:

- a makrostatistika és a HKF kategóriáinak megfeleltetése (amit közgazdaságilag el- térő jellegű kategóriák HKF-beli összevonása is megnehezít);<sup>12</sup>
- egyes tételeknek az NSZ-től eltérő besorolása a HKF-ben (például a borralaló és hálapénz munkajövedelemként és fogyasztásként való szerepeltetése);
- a háztartás-statisztikából hiányzó adatok (egyes termelési költségnemek, makrostatistikában is imputált tételek, biztosítások felosztása transzfer- és szolgáltatási díj részre az SNA szerint, közműfejlesztési hozzájárulás stb.) kezelése.

Természetesen, mint említettük, az adatbázis – az alkalmazás és a publikálás során ér- kező visszajelzéseket is hasznosítva – tovább finomítható. Köztudott azonban, hogy a „pontos” adatok délibábja kergetésének nincs értelme a feketegazdaság nagy aránya mi- att, amelyről aligha lehet teljes körű adatokhoz jutni.

## A SZÁMÍTÁSI EREDMÉNYEKRŐL

A számítások eredményeit a táblák részletesen mutatják be. Sok esetben a számok önmagukért beszélnek. Ennek ellenére az olvasó önálló értékelését és elemzéseit elő- segítendő, megpróbálok ízelítőt adni a sokirányú elemzési lehetőségekből.

Az elemzés jellege módszertanilag háromféle lehet: az egyes adatok önmagukban való értékelése, a jelen számítások eredményeinek egymással való összehasonlítása (belső vi- szonyítás) és végül más elemzések eredményeivel való összehasonlítás (külső viszonyí- tás).

Az adatok önmagukban való értékelése azt hiszem magától értetődő, hiszen egy ko- rábban nem ismert abszolút tényadat (például mennyi üzemanyagadót fizettek a gyerme- kes háztartások) az adott információt felhasználni kívánó döntéshozónak, illetve döntéselőkészítőnek nagyon fontos lehet.

<sup>12</sup> Például az illetékeknek és bírságoknak, az ingóságok és ingatlanok eladásából származó jövedelmeknek, a szociálpolitikai kedvezménynek és a vissza nem térítendő kölcsönöknek az összevont szerepeltetése, vagy a fogyasztási hitelek kamatainak a kapcsolódó termékre fordított kiadások részeként történő elszámolása.

A belső viszonyítás lehet az egyes kategóriák közötti (például adó/létszám, adó/jövedelem, egyenes adó/forgalmi adó) vagy rétegek közötti (abszolút összegek vagy fajlagos értékek aránya, illetve megoszlása).

A külső viszonyítás lehet térbeli és időbeli egyaránt. A „térbeli” összehasonlításon a külföldi szakirodalommal való összehasonlítást, az „időbelin” pedig korábbi hazai elemzések eredményeivel való összehasonlítást értünk.

A továbbiakban az egyes adatok önmagukban való (soknak vagy kevésnek) minősítésével nem foglalkozom.

A másik két összehasonlítási lehetőséget a következő alfejezetek mutatják be (a közismert tényekre, még ha fontosak is, általában nem térek ki, csak ha a számított érték váratlan vagy újszerű).

A számítási eredményeknek a (módszertani háttérből, részletességéből, az elvégzett korrekciókból, a feltüntetett helyi értékek számából stb. esetleg vélelmezhető) látszólagos nagyfokú pontossága<sup>13</sup> ellenére hangsúlyozni kell, hogy a HKF – a más országokban történt ilyen típusú adatfelvételekhez hasonlóan – nem tesz lehetővé pontos és vitathatatlan számítási eredményeket, de alkalmas lehet tárgyalási (továbbfejlesztési) alapnak és a hozzávetőleges fő arányok és tendenciák érzékeltetésére.

#### *A rétegek számított mutatóinak összehasonlítása*

Mindenekelőtt érdemes a rétegek néhány társadalmi–gazdasági állapotjellemzőjét szemügyre venni, amelyek a továbbiakban bemutatandó jövedelem- és adószerkezeteket is érthetőbbé tehetik.

Az olvasó saját elemzési lehetőségeit növelendő az 1. táblában az OECD ekvivalencia-skála szerinti fogyasztási egységeket is feltüntettem. A személygépkocsik valóságos (arány)száma ugyan nagyobb a tábla (a HKF) számainál, az egyes rétegek közötti arányokat azonban feltehetően jól tükrözi.

A 2. táblából látható az egyes rétegek kötődése az egyes szakágazatokhoz. A közismert (bár ezáltal számszerűsített és bizonyított) tényekre (mint például, hogy a szegények igen nagy arányban dolgoznak a mezőgazdaságban) nem térek ki. Érdekes viszont, hogy a középréteg nem átmenet a szegények és a jómódúak között, hanem mindkettőnél nagyobb arányban kötődik az oktatás és egészségügy ágazatokhoz. A kényszerűen aggregált adatokat tartalmazó táblából ugyan nem látható, de ugyanez a jelenség mutatkozott a gép-berendezésgyártás, az alapanyag-vegyipar, a bútorigar és a kultúra ágazatokban.

A 3. táblából látható, hogy a magasabb jövedelműek még nyugdíjból is többet kapnak mint a többiek. Hasonló a helyzet az ösztöndíjak esetében is, ami a főiskolás diákok miatt jelentkezik, jóllehet az 1. tábla szerint e csoportban az ösztöndíjklétszám nem lehet túl magas. E táblából látható az is, hogy mennyire zavaró a szociálpolitikai kedvezménynek a „vissza nem térítendő kölcsönökkel”<sup>14</sup> összevont szerepeltetése a HKF-ben. Ugyanis a táblából is érzékelhetően a két komponens feltehetően eltérő jövedelmi–demográfiai csoportoknál koncentrálódik.

<sup>13</sup> Ezek a szerző ellenőrizhetőségét és önellenőrzését segítik elő.

<sup>14</sup> Ezt a fából vaskarika szemléletet fejezte ki az a rendező is, aki azzal az indoklással nem adott kölcsön éppen *Eperjes Károly* színművészeknek, hogy „Mert maga visszaadja!”

1. tábla

## A háztartások helyzete és demográfiai jellemzőik

Kategória	Összes	Alacsony	Közepes	Magas	Városi	Községi	Gyermekes	Gyermektelen	Inaktív	Aktív	Ebből:					
		jövedelmű									gyermekes	gyermektelen	szegény	közép	jómódú	
		háztartás														
Létszám (ezer fő)	10051	4021	4023	2007	6303	3748	5498	4553	2892	7159	4834	2325	2867	2746	1546	
		Ezer főre jutó adatok														
Fogyasztási egységek száma	790	757	801	832	794	781	754	832	789	790	760	852	761	795	833	
20 évnél fiatalabb gyermek	235	342	187	115	233	237	429	0	108	286	423	0	380	265	147	
Alkalmazottak száma	301	222	309	443	319	270	300	302	14	417	340	577	310	450	555	
Vállalkozók száma	48	26	45	97	53	40	49	46	3	66	56	88	36	66	123	
Inaktív nyugdíjasok száma	277	186	365	281	268	291	68	529	721	97	51	193	84	108	102	
Segélyezetttek száma	73	129	43	20	64	88	102	38	82	69	83	41	112	51	22	
A lakás területe (ezer négyzetméter)	27	21	30	35	26	29	20	36	37	23	20	30	19	24	30	
Nyaráló (darab)	24	8	24	57	34	8	14	36	34	20	15	29	8	20	41	
Személygépkocsi összesen (darab)	139	83	143	244	140	137	128	153	94	157	139	196	98	167	249	

2. tábla

*A munkajövedelmek ágazati szerkezete*  
(százalék)

Szakágazat	Összes	Alacsony	Közepes	Magas	Városi	Községi	Gyermekes	Gyermektelen	Inaktív	Aktív	Ebből:					
		jövedelmű									háztartás	gyermekes	gyermektelen	szegény	közép	jómódú
1. Mező-hal-vadgazdálkodás	2,8	3,8	3,5	1,8	1,5	6,4	2,8	2,8	2,4	2,8	2,8	2,8	3,8	3,5	1,8	
2. Erdőgazdaság	0,3	0,3	0,5	0,2	0,2	0,6	0,4	0,3	1,4	0,3	0,3	0,2	0,3	0,5	0,1	
3. Bányászat	0,5	0,6	0,6	0,5	0,5	0,5	0,6	0,5	0,2	0,6	0,6	0,5	0,6	0,6	0,5	
4. Élelmiszer- és dohányipar	3,1	4,6	2,8	2,8	3,0	3,4	3,3	2,9	2,0	3,2	3,3	3,0	4,7	2,8	2,9	
5. Könnyűipar	3,5	5,3	3,7	2,7	3,5	3,6	3,8	3,3	1,9	3,6	3,8	3,4	5,5	3,8	2,7	
6. Vegyipar	2,9	2,5	3,0	3,0	3,0	2,7	2,7	3,2	1,3	3,0	2,7	3,3	2,6	3,1	3,0	
7. Építőanyag-ipar	0,9	0,7	0,7	1,1	0,3	2,4	0,8	1,0	0,3	0,9	0,8	1,0	0,7	0,7	1,1	
8. Kohászat	2,0	3,0	2,1	1,5	2,1	1,7	2,2	1,7	1,1	2,0	2,2	1,7	3,0	2,1	1,5	
9. Gépipar	5,3	5,3	5,3	5,3	5,0	6,2	5,2	5,5	5,8	5,3	5,1	5,5	5,3	5,2	5,3	
10. Egyéb feldolgozóipar	0,5	0,8	0,8	0,3	0,3	1,1	0,6	0,5	0,7	0,5	0,6	0,5	0,7	0,8	0,3	
11. Villamosenergia-, gáz-, hő-, vízellátás	2,1	1,9	1,9	2,4	2,0	2,6	2,2	2,0	0,9	2,2	2,3	2,1	1,9	2,0	2,4	
12. Építőipar	3,7	6,6	3,7	2,5	3,4	4,5	3,8	3,5	2,2	3,7	3,8	3,7	6,6	3,8	2,5	
13. Kereskedelem	10,1	11,3	9,7	10,0	11,2	7,0	9,4	10,9	6,0	10,3	9,4	11,3	11,5	9,8	10,1	
14. Vendéglátás	1,7	2,0	1,5	1,7	1,5	2,1	1,6	1,7	0,6	1,7	1,6	1,8	2,1	1,5	1,7	
15. Szállítás	4,6	5,3	5,4	3,7	4,3	5,5	5,3	3,9	2,2	4,7	5,3	4,0	5,3	5,6	3,8	
16. Távközlés	2,0	2,3	2,3	1,5	2,1	1,6	1,9	2,1	0,5	2,0	1,9	2,2	2,4	2,4	1,6	
17. Pénzügyi tevékenység	2,9	1,0	2,1	4,4	3,5	1,5	2,7	3,3	0,4	3,0	2,7	3,5	1,1	2,1	4,6	
18. Ingatlan-, gazdasági szolgáltatás	6,7	7,0	5,2	7,7	8,1	2,6	6,1	7,3	4,4	6,7	6,2	7,5	7,3	5,2	7,8	
19. Közigazgatás, közösségi szolgáltatás	9,9	8,4	9,4	11,0	10,0	9,9	9,4	10,7	7,8	10,0	9,4	10,9	8,4	9,6	11,0	
20. Oktatás	5,1	4,3	5,8	4,9	5,3	4,4	5,4	4,8	4,0	5,2	5,4	4,9	4,4	5,9	4,8	
21. Egészségügyi-szociális ellátás	4,5	5,4	5,7	3,2	4,7	3,9	4,4	4,7	6,8	4,4	4,4	4,5	5,5	5,7	2,9	
Keresetek összesen	75,3	82,4	75,7	72,2	75,6	74,4	74,4	76,3	52,7	76,1	74,5	78,2	83,7	76,9	72,4	
Mezőgazdasági kistermelés	5,1	6,6	6,0	3,6	2,8	11,5	4,3	6,0	40,1	3,7	4,0	3,4	5,1	4,2	2,8	
Vállalkozási jövedelem	19,7	11,0	18,3	24,2	21,6	14,1	21,3	17,7	7,2	20,2	21,5	18,4	11,2	18,9	24,8	
Munkajövedelem összesen (milliárd forint)	4 420	806	1 619	1 995	3 260	1 160	2 414	2 006	163	4 257	2 384	1 873	774	1 568	1 915	

3. tábla

## Egy főre jutó társadalmi és egyéb jövedelmek

Jövedelemtípusok (ezer forint/fő)	Összes	Alacsony	Közepes	Magas	Városi	Községi	Gyermekes	Gyermektelen	Inaktív	Aktív	Ebből:				
		jövedelmű									gyermekes	gyermektelen	szegény	közép	jómódú
		háztartás													
Táppénz	6,4	3,3	7,1	11,4	6,9	5,7	6,5	6,4	0,3	8,9	7,3	12,4	4,5	10,3	14,8
Nyugdíj, nyugdíjkiegészítő időskori járadék	93,3	47,6	120,4	130,5	97,0	87,0	20,7	180,9	247,5	31,0	15,9	62,4	21,8	35,1	40,8
Rendszeres segély	7,2	12,9	4,4	1,6	6,1	9,1	5,7	9,1	14,3	4,4	4,1	5,0	7,7	3,1	0,5
Nem rendszeres segély	0,8	0,8	1,0	0,5	0,8	0,8	0,6	1,1	1,7	0,4	0,4	0,5	0,6	0,3	0,2
Lakásfenntartási támogatás	0,4	0,5	0,4	0,1	0,5	0,2	0,3	0,5	0,8	0,2	0,3	0,1	0,4	0,2	0,0
Családi pótlék	12,2	19,5	9,1	4,0	11,6	13,3	21,8	0,6	7,0	14,3	20,9	0,6	21,0	12,7	4,9
Árvaellátás	1,9	1,9	1,9	2,1	2,2	1,6	3,2	0,4	2,3	1,8	2,5	0,4	1,1	2,2	2,4
Egyszeri anyasági támogatás	0,7	0,9	0,5	0,6	0,6	0,8	1,0	0,3	0,4	0,8	1,0	0,5	1,0	0,7	0,8
Gyed	0,1	0,1	0,1	0,1	0,1	0,1	0,2	0,0	0,0	0,1	0,2	0,1	0,2	0,1	0,2
Gyes	3,9	6,8	2,4	1,0	3,5	4,5	7,0	0,1	2,3	4,5	6,6	0,1	7,4	3,4	1,2
Gyet	0,9	1,9	0,3	0,1	0,8	1,3	1,7	0,0	0,9	0,9	1,4	0,0	1,8	0,5	0,2
Gyermekvédelmi pénzbeni támogatás	2,3	5,1	0,7	0,1	1,7	3,5	4,3	0,0	2,5	2,3	3,4	0,0	4,8	0,9	0,1
Gyermekvédelmi természetbeni támogatás	0,3	0,6	0,2	0,1	0,3	0,4	0,6	0,0	0,4	0,3	0,4	0,0	0,5	0,2	0,1
Ösztöndíj	1,2	0,8	1,0	2,5	1,6	0,6	1,0	1,4	1,1	1,3	1,0	2,0	0,5	1,4	2,7
Munkanélküli járadék	4,7	6,3	4,2	2,4	4,5	5,0	4,8	4,6	4,2	4,9	4,4	5,9	6,2	4,6	2,9
Munkanélküli jövedelempótló támogatás	2,4	5,0	1,0	0,3	2,1	3,0	2,8	2,0	3,9	1,8	1,8	1,9	3,4	1,0	0,3
Ápolási díj	0,3	0,5	0,3	0,1	0,3	0,3	0,3	0,3	0,5	0,3	0,3	0,2	0,4	0,2	0,0
Külföldről származó társadalmi juttatás	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Természetben kapott jövedelem (étkezés, szociális ellátás)	1,0	1,0	0,8	1,4	1,0	1,0	0,8	1,2	1,4	0,8	0,7	1,0	0,8	0,6	1,3
Szociálpolitikai kedvezmény, vissza nem fizetendő támogatás	2,8	5,5	0,3	2,4	1,0	6,0	4,6	0,7	5,2	1,9	2,4	0,8	3,0	0,2	3,0
Pénzbeni és hasonló társadalmi juttatás összesen	143,1	121,2	156,1	161,3	142,6	144,0	88,0	209,7	296,8	81,1	74,9	94,0	87,0	77,7	76,2
Felvett kamat, osztalék	0,6	0,1	0,5	1,7	0,6	0,6	0,3	1,0	0,6	0,6	0,3	1,2	0,1	0,4	1,8
Egyéb jövedelem	3,8	1,9	3,2	8,5	3,6	4,0	2,8	4,9	4,2	3,6	2,8	5,2	2,0	3,1	7,4
Társadalmi és egyéb jövedelem összesen	147,5	123,2	159,8	171,5	146,8	148,7	91,1	215,6	301,5	85,3	78,0	100,3	89,1	81,2	85,4



4. tábla

## Az adóterhelés

Kategória	Összes	Alacsony	Közepes	Magas	Városi	Községi	Gyermekes	Gyermektelen	Inaktív	Aktív	Ebből:					
		jövedelmű									háztartás	gyermekes	gyermektelen	szegény	közép	jómódú
Egy főre (ezer forint)																
Nyugdíj-, egészségügyi és munkavállalói járulékok	25,2	13,0	23,5	52,9	28,5	19,5	26,1	24,1	1,4	34,8	29,2	46,3	17,6	34,1	67,8	
Személyi jövedelemadó	65,3	22,9	56,4	168,3	79,2	42,0	65,9	64,7	5,5	89,5	74,3	121,1	31,6	81,5	211,1	
Illeték, bírság	2,1	0,9	1,4	6,1	2,5	1,5	1,9	2,4	1,7	2,3	2,2	2,6	1,0	1,5	6,4	
Járműadó, -illeték	1,0	0,6	0,9	1,7	1,0	0,8	0,9	1,0	0,6	1,1	1,0	1,3	0,7	1,1	1,8	
Építményadó és egyéb helyi adó	1,9	1,2	2,1	3,1	2,0	1,8	1,3	2,6	2,6	1,6	1,4	2,0	1,1	1,6	2,4	
Lakásberuházáson nettó áfa	2,5	1,3	2,1	5,7	2,4	2,7	2,2	2,8	1,5	2,9	2,4	3,9	1,5	2,4	6,4	
Vám	3,9	2,5	3,6	7,3	4,1	3,6	3,6	4,3	3,2	4,2	3,7	5,2	2,6	3,8	7,8	
Belföldi forgalmi adók	84,1	56,8	81,8	143,1	88,4	76,8	75,1	94,9	71,2	89,2	78,8	111,0	59,6	85,6	150,7	
Összes adó és elvonás	186,0	99,2	171,8	388,2	208,1	148,7	177,0	196,8	87,8	225,6	193,0	293,5	115,7	211,6	454,4	
Belföldi forgalmi adókból:																
dohánytermékek	10,0	9,4	10,0	11,2	10,2	9,7	8,8	11,4	9,2	10,3	8,8	13,3	9,4	10,6	11,5	
szeszes italok	5,0	3,5	4,4	9,0	5,3	4,5	3,9	6,2	4,0	5,3	4,2	7,6	4,0	4,6	9,1	
köolajfeldolgozási termékek	12,7	7,7	12,4	23,2	11,8	14,2	11,6	14,0	8,6	14,3	12,5	18,2	8,5	14,3	25,1	
Egy főre jutó összes jövedelem	587,3	323,7	562,2	1165,6	664,0	458,1	530,3	656,1	358,0	679,9	571,2	905,7	358,8	652,3	1324,2	
Adóterhelés (adó/jövedelem) (százalék)	31,66	30,63	30,57	33,30	31,34	32,46	33,37	30,00	24,52	33,19	33,78	32,40	32,24	32,44	34,31	
Egyenes adó/munkajövedelem (százalék)	27,34	17,91	19,85	22,25	20,83	19,88	20,93	20,16	12,14	20,90	20,99	20,79	18,24	20,25	22,51	
Millió forint																
Nyugdíj-, egészségügyi és munkavállalói járulékok	252 967	52 226	94 645	106 096	179 840	73 127	143 151	109 816	3 915	249 052	141 296	107 756	50 586	93 650	104 817	
Személyi jövedelemadó	656 624	92 187	226 673	337 764	499 165	157 459	362 214	294 410	15 900	640 724	359 144	281 580	90 494	223 853	326 377	
Illeték, bírság	21 557	3 709	5 573	12 275	15 933	5 624	10 690	10 867	4 832	16 725	10 578	6 148	2 825	4 059	9 841	
Járműadó, -illeték	9 451	2 265	3 702	3 484	6 597	2 854	4 838	4 613	1 692	7 759	4 641	3 118	1 968	3 020	2 771	
Építményadó és egyéb helyi adó	19 215	4 675	8 366	6 174	12 400	6 815	7 143	12 072	7 893	11 322	6 604	4 718	3 073	4 500	3 750	
Lakásberuházáson nettó áfa	25 000	5 171	8 352	11 477	14 914	10 086	12 245	12 755	4 260	20 740	11 751	8 989	4 307	6 488	9 944	
Vám	39 223	9 949	14 669	14 605	25 731	13 492	19 539	19 684	9 277	29 946	17 948	11 998	7 357	10 506	12 083	
Belföldi forgalmi adók	844 919	228 435	329 225	287 260	556 995	287 924	413 072	431 847	206 031	638 887	380 898	257 989	171 029	234 934	232 924	
Összes adó és elvonás	1 868 956	398 616	691 205	779 134	1 311 573	557 382	972 891	896 064	253 800	1 615 156	932 860	682 296	331 638	581 010	702 508	
Belföldi forgalmi adókból:																
dohánytermékek	100 275	37 674	40 194	22 407	64 082	36 193	48 565	51 710	26 536	73 738	42 722	31 016	26 961	29 010	17 767	
szeszes italok	49 934	14 169	17 622	18 142	33 138	16 796	21 664	28 269	11 663	38 271	20 514	17 756	11 426	12 711	14 134	
köolajfeldolgozási termékek	127 523	30 946	49 975	46 602	74 264	53 259	63 938	63 585	24 974	102 549	60 337	42 211	24 410	39 310	38 829	
Összes jövedelem (házipírtési, imputált és ttj.nélk.)	5 902 350	1 301 443	2 261 305	2 339 601	4 185 355	1 716 995	2 915 594	2 986 756	1 035 244	4 867 106	2 761 242	2 105 864	1 028 707	1 791 053	2 047 346	

5. tábla

## Az adóterhelés adófajtánkénti szerkezete

Kategória	Összes	Alacsony	Közepes	Magas	Városi	Községi	Gyermekes	Gyermektelen	Inaktív	Aktív	Ebből:				
		jövedelmű									gyermekes	gyermektelen	szegény	közép	jómódú
		háztartás													
<b>A rétegek által fizetett összes adó százalékában</b>															
Nyugdíjjárulék és munkavállalói járulék	13,6	13,1	13,7	13,6	13,7	13,1	14,7	12,3	1,5	15,4	15,2	15,8	15,3	16,1	14,9
Személyi jövedelemadó	35,1	23,1	32,8	43,3	38,1	28,3	37,2	32,9	6,3	39,7	38,5	41,2	27,3	38,5	46,5
Illeték, bírság	1,2	0,9	0,8	1,6	1,2	1,0	1,1	1,2	1,9	1,0	1,1	0,9	0,9	0,7	1,4
Járműadó, -illeték	0,5	0,6	0,5	0,4	0,5	0,5	0,5	0,5	0,7	0,5	0,5	0,5	0,6	0,5	0,4
Építményadó és egyéb helyi adó	1,0	1,2	1,2	0,8	0,9	1,2	0,7	1,3	3,1	0,7	0,7	0,7	0,9	0,8	0,5
Lakásberuházáson nettó áfa	1,3	1,3	1,2	1,5	1,1	1,8	1,3	1,4	1,7	1,3	1,3	1,3	1,3	1,1	1,4
Vám	2,1	2,5	2,1	1,9	2,0	2,4	2,0	2,2	3,6	1,8	1,9	1,8	2,2	1,8	1,7
Belföldi forgalmi adók	45,2	57,3	47,7	36,9	42,5	51,7	42,5	48,2	81,2	39,6	40,8	37,8	51,6	40,5	33,2
Összes adó és elvonás	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
Belföldi forgalmi adókból:															
dohánytermékek	5,37	9,45	5,82	2,88	4,89	6,49	4,99	5,77	10,46	4,57	4,58	4,55	8,13	4,99	2,53
szeszes italok	2,67	3,55	2,55	2,33	2,53	3,01	2,23	3,15	4,60	2,37	2,20	2,60	3,45	2,19	2,01
kőolajfeldolgozási termékek	6,82	7,76	7,23	5,98	5,66	9,56	6,57	7,10	9,84	6,35	6,47	6,19	7,36	6,77	5,53
<b>A rétegek összes jövedelme százalékában</b>															
Nyugdíjjárulék és munkavállalói járulék	4,29	4,01	4,19	4,54	4,30	4,26	4,91	3,68	0,38	5,12	5,12	5,12	4,92	5,23	5,12
Személyi jövedelemadó	11,12	7,08	10,02	14,44	11,92	9,17	12,42	9,86	1,54	13,16	13,00	13,37	8,80	12,50	15,94
Illeték, bírság	0,37	0,29	0,25	0,52	0,38	0,33	0,37	0,36	0,47	0,34	0,38	0,29	0,27	0,23	0,48
Járműadó, -illeték	0,16	0,17	0,16	0,15	0,16	0,16	0,17	0,15	0,16	0,16	0,17	0,15	0,19	0,17	0,14
Építményadó és egyéb helyi adó	0,33	0,36	0,37	0,26	0,30	0,40	0,24	0,40	0,76	0,23	0,24	0,22	0,30	0,25	0,18
Lakásberuházáson nettó áfa	0,42	0,40	0,37	0,49	0,36	0,59	0,42	0,43	0,41	0,43	0,43	0,43	0,42	0,36	0,49
Vám	0,66	0,77	0,65	0,62	0,61	0,79	0,67	0,66	0,90	0,62	0,65	0,57	0,72	0,59	0,59
Belföldi forgalmi adók	14,31	17,55	14,56	12,28	13,31	16,76	14,17	14,46	19,90	13,13	13,79	12,25	16,62	13,11	11,37
Összes adó és elvonás	31,66	30,63	30,57	33,30	31,34	32,46	33,37	30,00	24,52	33,19	33,78	32,40	32,24	32,44	34,31
Belföldi forgalmi adókból:															
dohánytermékek	1,70	2,89	1,78	0,96	1,53	2,11	1,67	1,73	2,56	1,52	1,55	1,47	2,62	1,62	0,87
szeszes italok	0,85	1,09	0,78	0,78	0,79	0,98	0,74	0,95	1,13	0,79	0,74	0,84	1,11	0,71	0,69
kőolajfeldolgozási termékek	2,16	2,38	2,21	1,99	1,77	3,10	2,19	2,13	2,41	2,11	2,19	2,00	2,37	2,19	1,90

A 4. tábla szerint a jómódúak egy főre jutó jövedelme 3,6-szerese a szegényekének. Ebből adót is majdnem 4-szer többet fizetnek fejenként. Következésképpen a százalékos adóteher enyhén progresszív. Összességében vett progresszív adózásról tehát az előző kormányzat utolsó évében nemigen beszélhetünk. Ez lényegében igaz a munkajövedelmek adóterhelésére is (nyilván a munkaadókat terhelő lineáris elvonások, sőt a degresszív egészségügyi hozzájárulás miatt ez általánosabban is igaz). A szegények magas adóterhelését jelentős mértékben a forgalmi adók okozták (a szociálisan érzéketlen, és rendkívül magas adómértékek miatt, amin egyes későbbi intézkedések, mint például a tankönyvek áfa-mentessé tétele igyekeztek enyhíteni).

Figyelemreméltó, hogy még az aktív háztartásokon belül is a gyermekes családok százalékos adóterhelése magasabb, mint a náluk átlagosan másfélszer több fejenkénti jövedelmet élvező gyermekteleneké. Ez egyértelműen mutatja, hogy mennyire szükség volt a gyermekek után járó jelentős adókedvezmény bevezetésére<sup>15</sup> a családi adózás megoldatlansága (ami miatt a családfenntartó nemcsak a marginális kulccsal kell adózzon a gyakran létminimumot sem elérő egy főre jutó jövedelmek után, hanem ekkor még az adójóváírástól is megfosztják), és a családi pótlék elértéktelenedése miatt.

Az 5. táblából jól látható, hogy a háztartásokat terhelő adók milyen nagy mértékben koncentrálnak a személyi jövedelemadó-járulékok-belföldi forgalmi adók hármásra, míg a helyi (vagyon- stb.) adók, vámok és illetékek aránya nemzetközileg és a régebbi (1935. évi) magyar arányokhoz képest is milyen alacsony. Ráadásul a szegények jövedelmüknek még jóval nagyobb hányadát is fizetik be forgalmi adóként, mint a többiek. Ez még akkor is így van, ha az élvezeti cikkek forgalmi adóját (amelyek fogyasztása nem lét-szükséglet) leválasztjuk az összes forgalmi adóból. Sajnos e tétel nagyobb arányú a szegényeknél, ami rámutat egyfajta csapdahelyzetre, illetve az állam ezzel kapcsolatos teendőinek szociális fontosságára is.<sup>16</sup>

#### *Összehasonlítás más elemzésekkel*

A 6–8. táblák néhány más elemzés eredményével hasonlítják össze a rétegek jövedelem-szintjére és adóterheire vonatkozó becsléseinket. Az összehasonlítást gátló számos módszertani problémára ezúttal nem térek ki (bruttó vagy nettó jövedelem, adók köre, csoportosítás egy főre, vagy egy fogyasztási egységre jutó jövedelem vagy fogyasztás alapján, korrigált vagy nyers adatok stb.), csupán annyit jegyzek meg, hogy mivel más források többnyire jövedelmi deciliseket (népességtizedeket) mutatnak be, ezek közvetlenül összehasonlíthatók tanulmányunk jövedelemkategóriáival. A Matolcsy-féle rétegekkel való összehasonlításhoz kisebb mértékű átcsoportosítást kellett tennem az ott szereplő „szerénysorú” népességből a magasabb jövedelműek felé. Ezt konkrétan úgy oldottam meg, hogy a mintegy 81,2 százalékos létszámú szegény rétegből a 80 százalékos feletti részt leválasztottam azzal a feltevessel, hogy ennek az 1,2 százaléknak az egy főre jutó jövedelme és adóterhe a szegény és jómódú rétegek fajlagos értékeinek átlaga. Hasonló módon korrigáltam a „nagyjövedelműek” csoportját is a népesség 20 százalékára (felső kvintilisre).

<sup>15</sup> Az más kérdés, hogy a más adókedvezményekhez és családtámogatási formákhoz hasonlóan ennek sem garantált az értékmegőrzése, noha a gyermekvállalás hosszú távú elkötelezettséget kíván.

<sup>16</sup> A forgalmi adókra hasonló eredményeket kapott egy korábbi tanulmány szerzője is (*Keszthelyiné-Kéki-Pintye*; 1995), de az alsó áfa-kulcs időközbeni emelése miatt azóta a helyzet romlott.

6. tábla

*A jövedelem- és adómegoszlás összehasonlítása*  
(százalék)

A felmérés megnevezése	Összesen	Alsó 40 százalék	Közepes jövedelműek	Felső 20 százalék
Jövedelmek megoszlása (elől a vizsgálat tárgyéve):				
1935 – <i>Matolcsy</i> (1938) (csoportok kvintilisekre átszámítva)	100,0		43,3	56,7
1987 – Háztartásstatisztika (nettó, KSH) ( <i>Ecostat</i> ; 2001)	100,0	25,1	40,2	34,7
1989 – Incidencia vizsgálat (bruttó, KSH) ( <i>Fajth-Kupa</i> )	100,0	24,7	39,7	35,6
1989 – Incidencia vizsgálat (nettó, KSH)	100,0	25,7	39,8	34,5
1991 – Háztartásstatisztika (nettó, KSH, Unicefnek küldött)	100,0	27,3	40,6	32,1
1991 – Háztartásstatisztika (bruttó, KSH, Unicefnek küldött)	100,0	26,1	39,5	34,4
1993 – Háztartásstatisztika eredeti ( <i>Szabóné</i> ; 1996)	100,0	25,8	40,8	33,4
1993 – Háztartásstatisztika korrigált ( <i>Szabóné</i> ; 1996)	100,0	24,4	40,7	34,9
1993 – TÁRKI	100,0	22,3	38,6	39,1
1995 – Háztartásstatisztika (nettó, KSH) ( <i>Ecostat</i> ; 2001)	100,0	21,7	39,2	39,1
1996 – TÁRKI	100,0	20,0	39,4	40,6
1997 – Háztartásstatisztika (nettó, KSH) ( <i>Ecostat</i> ; 2001)	100,0	20,2	38,5	41,3
1998 – <i>Jelen tanulmány adataiból számított</i>	100,0	22,0	38,3	39,7
1999 – Háztartásstatisztika (nettó, KSH) ( <i>Ecostat</i> ; 2001)	100,0	20,0	38,0	42,0
Összes adó és elvonás				
1935 – <i>Matolcsy</i> (1938) (csoportok kvintilisekre átszámítva)	100,0		37,0	63,0
1991 – Háztartásstatisztika ( <i>Newbery-Révész</i> ; 1997)	100,0	20,3	36,0	43,7
1998 – <i>Jelen tanulmány adataiból számított</i>	100,0	21,3	37,0	41,7
Egyes adónemek				
1935 – Egyenes adók, <i>Matolcsy</i> (1938) (kvintilisekre)	100,0		23,4	76,6
1989 – Egyenes adók, Incidencia vizsgálat (KSH)	100,0	19,3	39,0	41,7
1989 – Szja, Incidencia vizsgálat (KSH)	100,0	15,3	37,4	47,3
1991 – Szja ( <i>Newbery-Révész</i> ; 1997)	100,0	14,3	32,0	53,7
1998 – Szja, <i>jelen tanulmány adataiból számított</i>	100,0	14,0	34,5	51,5
1989 – Nyugdíjjárulék, Incidencia vizsgálat (KSH)	100,0	22,2	41,1	36,7
1991 – Nyugdíjjárulék ( <i>Newbery-Révész</i> ; 1997)	100,0	20,5	38,5	41,0
1998 – Nyugdíjjárulék, <i>jelen tanulmány adataiból számított</i>	100,0	20,6	37,4	42,0
1935 – Indirekt adók vám nélkül, <i>Matolcsy</i> (1938) (kvintilisre)	100,0		52,5	47,5
1991 – Indirekt adók vám nélkül ( <i>Newbery-Révész</i> ; 1997)	100,0	27,2	39,2	33,6
1998 – Indirekt adók vám nélkül, <i>jelen tanulmányból</i>	100,0	27,0	38,9	34,1

7. tábla

*Az összesített adóterhelés összehasonlítása a jövedelem százalékában*  
(indirekt adókkal)

A felmérés megnevezése	Összesen	Alsó 40 százalék	Közepes jövedelműek	Felső 20 százalék
1935 – <i>Matolcsy</i> (1938) (kvintilisekre)	25,1		21,4	27,9
1991 – Háztartás statisztika ( <i>Newbery-Révész</i> ; 1997)	26,9	20,9	24,5	34,2
1998 – <i>Jelen tanulmány adataiból számított</i>	31,7	30,6	30,6	33,3

A számszerű eredményekből a következők emelhetők ki.

– Az 1935. évi állapothoz képest nőtt az alsó két jövedelmi csoport részesedése, az adórendszer még 1991-ben is megfigyelhető (a személyi jövedelemadó bevezetésekor rövid ideig elég nagy) progresszió azonban lényegesen csökkent (bár már *Matolcsy* is csökkenő adóterhelési arányt talált a legfelső rétegnél). Ezt egy korábbi cikkemben (*Newbery-Révész*; 2000) részletes és szemléletes számításokkal mutattam be.

– A jövedelemegyenlőtlenség 1991-ig csökkent, utána ismét nőtt. Ennek nyilvánvalóan a munkanélküliség hirtelen felszökése lehetett az egyik fő oka, de a kérdés számszerűen csak a jövedelemszerkezetek időbeli összehasonlításával elemezhető, ami egy részletesebb tanulmányt kíván.

– A nagyjövedelműek relatív pozíciója a másik két réteghez képest a 90-es években, ha nem is drámai mértékben és kisebb-nagyobb (feltehetően jórészt csak statisztikai mérési) ingadozásokkal, de folyamatosan javulni látszik.

8. tábla

*A jövedelemmegoszlás nemzetközi összehasonlítása*  
(százalék)

A felmérés megnevezése	Összesen	Alsó 40 százalék	Közepes jövedelműek	Felső 20 százalék
Jövedelem nagysága szerinti csoportok jövedelmi részesedése				
1998 – Magyarország (Jelen tanulmányból) B	100,0	22,0	38,3	39,7
Átalakuló gazdaságok				
1996 – Szlovákia B	100,0	27,7	41,0	31,3
1996 – Csehország B	100,0	24,8	39,4	35,8
1994 – Románia B	100,0	22,5	40,2	37,3
1997 – Bulgária B	100,0	24,0	39,3	36,7
Európai nagy, illetve hozzájuk hasonló méretű országok				
1994 – Németország B	100,0	21,4	40,2	38,4
1995 – Franciaország B	100,0	19,8	40,0	40,2
1993 – Görögország B	100,0	19,9	39,7	40,4
1994/5 – Portugália B	100,0	18,9	37,7	43,4
1992 – Svédország B	100,0	24,1	41,3	34,6
Európán kívüli nagy országok				
1997 – Egyesült Államok B	100,0	15,7	38,0	46,3
1997 – Brazília B	100,0	8,3	28,8	62,9
1998 – Kína B	100,0	16,1	37,3	46,6
Fogyasztás nagysága szerinti csoportok fogyasztási részesedése				
1998 – Magyarország A	100,0	24,7	41,0	34,3
1998 – Lengyelország A	100,0	20,6	39,7	39,7
1999 – Ukrajna A	100,0	22,1	40,1	37,8
1999 – Oroszország A	100,0	13,0	33,4	53,6
1997 – India A	100,0	19,7	34,3	46,0
Régebbi adatok (a tendenciák érzékeltetésére):				
1985 – USA B	100,0	15,7	42,4	41,9
1988 – Franciaország B	100,0	17,4	40,7	41,9

*Megjegyzés.* A – fogyasztási szint csoportok fogyasztási megoszlása. B – egy főre jutó jövedelem szerinti csoportok jövedelemmegoszlása.

*Forrás:* World Development Reports 1997, 2001.

Az egyes felmérések adatai egyébként jól egybevágóak, ami különböző módszertani megközelítéseiket, értelmezéseiket és gyakorlati lebonyolítási technikáikat tekintve meglepő, de növeli az egyes módszerek valószínűsíthető megbízhatóságát. Ez is rámutat, hogy többféle megközelítés nem rontja egymás hitelét, hanem erősíti, feltéve, ha szakmailag korrektilt végzik az elemzést.

A kapott jövedelemeloszlási mutatók nemzetközi összehasonlításánál a *World Development Report* (1997, 2001) kiadványaira támaszkodtam. Mivel ez utóbbi a számítógépes világhálón bárki számára hozzáférhető ([www.worldbank.org/data](http://www.worldbank.org/data)) – rövid módszertani leírással együtt, amiből ugyan csak valószínűsíthető, hogy a bruttó jövedelmekkel dolgoztak, de „minden lehetőt megtettek” az összehasonlíthatóság biztosítása érdekében – , a mintegy másfélszáz országra közölt adatokból csak néhány, számunkra meghatározó viszonyítási alapként szolgáló ország adatait mutatom be. A szomszédos és sok tekintetben követni kívánt Ausztriára sajnos nem volt friss adat (csak 1985. évi), Németország időbeli összehasonlítása pedig a német újraegyesítés miatt nem lehetséges korrekt módon. Ezért ez utóbbinál a francia változásokat mutatom be. A felsorolt országok között a szomszédos, az átalakuló, a nagy és hozzánk hasonló méretű európai unióbeli tagállamok, valamint a világ nagy államai szerepelnek. Az adók megoszlására egyelőre nem jutottam elég széles körű, és módszertanilag megfelelő (világos tartalmú) adatokhoz. Ez további kutatás tárgya lehet.

Az országok adataiból úgy tűnik, hogy a magyar jövedelemegyenlőtlenség Európában nem haladja meg az átlagot, a középosztály jövedelemrészesevé válása viszont alacsony. Az Európai Unióban csak Portugáliában kisebb a középosztály jövedelemrészesevé válása, mint nálunk. A más földrészekre jellemző jövedelemegyenlőtlenségtől szerencsére határozottan távol állunk. Igaz a nyugat-európai átlagos részesevé válástól csak mintegy 2 százalékpont az elmaradás, de az Egyesült Államokéval egyező mérték elgondolkodtató (senki se kétli, hogy az Egyesült Államokban a helyzet rosszabb, mint Nyugat-Európában), és hogy mindössze további 4 százalékpont már Oroszország és India szintjére süllyesztene minket.

Franciaország és az Egyesült Államok adataiból nem szűrhetők le egyértelmű tendenciák (a két Világbank jelentés sok országának együttes elemzése egy másik tanulmányt érdemelne). A legszegényebbek (a táblából nem látható alsó 20 százalék) részesevé válása nőtt mindkét országban a második kvintilis rovására. Ha ezt a második kvintilist a középosztályhoz számítjuk, akkor a francia és az egyesült államokbeli 3-4. kvintilisre vonatkozó aránycsökkenés alapján feltételezhetjük, hogy a középosztály súlya csökkent (ez a frusztráció mutatkozhatott meg az időközben ott lezajlott jobboldali előretérésben). (Lásd a 8. táblát.)

\*

Az elemzés már a jelenlegi fázisában érdekes eredményekre vezetett. A részletes dokumentáció és a téma időszerűsége révén remélhető, hogy a szakmától számos hasznos visszajelzés érkezik, ami az alkalmazott módszer érdemi továbbfejlesztését teszi lehetővé. Érdemes lenne más kutatóműhelyek és kormányzati szervek bevonásával kiterjeszteni és elmélyíteni a vizsgálatokat, különösen a támogatások, adókedvezmények beépítésével és részletesebb felbontások (belföldi termékadófajták, régiók stb.) alkalmazásával, valamint más adatbázisokkal (például az időmérleg-felvétel, vagy a személyijövedelemadó-bevallások adataival) való összekapcsolással. Különösen az adómegoszlásra is kiterjedő nemzetközi összehasonlítások tennék gazdagabbá az elemzést. Az OECD, a Világbank, az Európai Unió anyagai, valamint az átalakuló gazdaságokra vonatkozó információk alapján helyzetünk és lehetőségeink sokkal jobban kirajzolódhatnak.

9. tábla

*Jövedelmek és adók összefoglaló táblája*  
(milliárd forint)

Kategória	Összesen	Alacsony	Közepes	Magas	Városi	Községi	Gyermekes	Gyermektelen	Inaktív	Aktív	Ebből:					
		jövedelmű									háztartás	gyermekes	gyermektelen	szegény	közép	jómódú
Munka és egyéb termelési jövedelem	4 420	806	1 619	1 995	3 260	1 160	2 414	2 006	163	4 257	2 384	1 873	774	1 568	1 915	
Társadalmi jövedelem	1 439	487	628	324	899	540	484	954	858	580	362	218	249	213	118	
Egyéb jövedelmek	44	8	15	21	26	17	17	27	14	30	15	15	6	10	14	
Bruttó jövedelem összesen	5 902	1 301	2 261	2 340	4 185	1 717	2 915	2 987	1 035	4 867	2 761	2 106	1 029	1 791	2 047	
Egyenes adók	960	155	339	466	714	246	528	432	34	926	522	403	149	329	448	
Belföldi forgalmi adók, vám	909	244	352	313	598	312	445	464	220	690	411	279	183	252	255	
Összes adóelvonás	1 869	399	691	779	1 312	557	973	896	254	1 615	933	682	332	581	702	
Nettó jövedelem	4 033	903	1 570	1 560	2 873	1 160	1 942	2 091	781	3 252	1 828	1 424	697	1 210	1 345	

*Megjegyzés.* A kerekítések következtében az egyes részösszegek nem adják ki az összesen oszlopban található értéket.

Ha a Központi Statisztikai Hivatal kidolgozza a mai magyar háztartásokra érvényesnek tekinthető ún. „fogyasztási egységekre” átszámító kulcsokat, akkor az egy fogyasztási egységre vetített mutatókkal is elvégezhető lesz a számítás.

Egy másik fontos továbblépésként érdemes lenne a modellháztartások bevezetőben említett módszerével megkísérelni a jelen tanulmány eredményeinek aktualizálását, azaz az 1998. évre kapott eredményekből kiindulva a jelenlegi kormányzati ciklus (1998–2002) alatti változások számszerűsítését.

#### IRODALOM- ÉS FORRÁSJEGYZÉK

- ADELMAN, I. – BERCK, P. – VUJOVIC, D. (1991): Using social accounting matrices to account for distortions in non-market economies. *Economic Systems Research*, 3. évf. 3. sz.
- Ágazati Kapcsolatok Mérlege. (2001a) Központi Statisztikai Hivatal, Budapest.
- Családi költségvetés 1998. Adattár. (1999b) Központi Statisztikai Hivatal, Budapest.
- ECOSTAT (2001): *A gazdaság fejlődése 2001-ben – Makrogazdaság és vállalkozások*. Az Ecostat Gazdaságelemző és Informatikai Intézet „MONITOR 2001/I. negyedév” c. kiadványa.
- FAJTH G. – KUPA M. (1990): Incidence Study '90: The Hungarian social policy: system and distribution of incomes of households. A világbank részére készített tanulmány. Központi Statisztikai Hivatal, Budapest.
- GÁL R. – SIMONOVITS A. – SZABÓ M. – TARCALI G. (2000): *A korosztályi elszámolás Magyarországon*. TÁRKI Társadalomtudományi Tanulmányok 22.
- GFK HUNGÁRIA (1999): *Phare turizmusfejlesztési program – turisztikai felmérés*. Kézirat.
- HORVÁTH E. (1999): *A nemzetközi aktív turizmus multiplikátor hatásainak becslése input-output modell alkalmazásával*. PhD-értekezés, Budapesti Közgazdaságtudományi és Államigazgatási Egyetem. Rövidített változata megjelent a Turizmus Bulletin 1999 márciusi számában „A Magyarországra látogató külföldiek fogyasztási szokásai” címmel.
- HÜTTL A. (2000): *A turizmus hatása a jövedelmekre, a foglalkoztatásra és a költségvetés helyzetére. Magyarország turizmus szatellit számlája – megvalósíthatósági tanulmány*. Kézirat.
- KESZTHELYINÉ RÉDEI M. – KÉKI ZS. – PINTYE J. (1995): *Az indirekt adók újraelosztási hatásai (Számítások az 1993. évi háztartási költségvetési felvétel alapján)*. Kézirat.
- Külföldiek fogyasztási szokásai, 1994. (1996) Központi Statisztikai Hivatal, Budapest.
- Lakásstatisztikai évkönyv 1998. (1999d) Központi Statisztikai Hivatal, Budapest.
- LECOMBER, J. R. C. (1975): A critique of methods of adjusting, updating, and projecting matrices. In: *Estimating and projecting input-output coefficients*. Input-Output Publishing Company, London. 1–25. old.
- Létminimum, 2000. (2001b) Központi Statisztikai Hivatal, Budapest.
- A Magyar Köztársaság 1998. évi költségvetésének zárszámadásáról. (1999) Pénzügyminisztérium, Budapest.
- Magyarország Nemzeti Számlái 1996–1998. (2000) Központi Statisztikai Hivatal, Budapest.
- Magyarország Nemzeti Számlái 1998–1999. (2001c) Központi Statisztikai Hivatal, Budapest.
- Magyarország – Szegénység és szociális támogatások. (1996) Világbank.
- MATOLCSY M. (1938): *A magyarországi jövedelem- és adóteher-megoszlás*. Magyar Gazdaságkutató Intézet, Budapest.
- NEWBERY, D. – RÉVÉSZ, T. (1997): *The burden and disincentive effect of the Hungarian personal tax system*. CEPR Discussion Paper, No. 1570. London.
- NEWBERY, D. – RÉVÉSZ, T. (2000): The evolution of the tax structure of a reforming transitional economy: Hungary 1988–98. *International Tax and Public Finance*, 7. sz. 209–240. old.
- RÉVÉSZ T. – ZALAI E. (1992): Adók és támogatások az átalakuló magyar gazdaságban. *AULA. Társadalom és gazdaság*, 2. sz. 30–54. old.
- RÉVÉSZ T. – ZALAI E. (1993): An analysis of the economic system of Hungary within a SAM (Social Accounting Matrix) framework. In: COHEN, S. I. (szerk.): *Patterns of economic restructuring for Eastern Europe*, Avebury-Aldershot.
- RÉVÉSZ T. (1994a): How the Hungarian households fared during the first period of transition. *Economics of Transition*, 2. évf. 1. sz. 95–101. old.
- RÉVÉSZ T. (1994b): *An analysis of the representativity of the Hungarian Household Budget Survey Samples*. Discussion Paper on Economic Transition, No. DPET 9403, University of Cambridge.
- RÉVÉSZ T. (1995): Háztartásstatisztika – Érvényességvizsgálat. *Statisztikai Szemle*, 73. évf. 1. sz. 31–49. old.
- RÉVÉSZ T. (1996): *Analysis of changes in the stratas' income and consumption by reconciling macroeconomic and survey data*, „Applied macro and micro economic modelling for European and Former Soviet transition economies” (Kézirat.)
- RÉVÉSZ T. (1997): *The impact of changes in the household related tax and benefit system analysed by the HUMUS Computable General equilibrium Model*. Az European Commission „Cooperation in Science and Technology” (ERB-CIPA-CT93-0230) kutatási projekt keretében a Budapesti Közgazdaságtudományi Egyetemen készített tanulmány. (Kézirat.)
- RÉVÉSZ T. (2001): A turizmus költséghatás-elemzése SAM-moddellel. *Statisztikai Szemle*, 79. évf., 10–11. sz. 825–847. old.
- Statisztikai évkönyv 1998. (1999a) Központi Statisztikai Hivatal, Budapest.
- SZABÓ SNÉ. (1996): Becslés a „valódi” jövedelemszintre és -szóródásra. *Statisztikai Szemle*, 74. évf. 2. sz. 126–124. old.
- Szociális statisztikai évkönyv 1998. (1999c) Központi Statisztikai Hivatal, Budapest.
- TÓTH I. J. (1997): *Az adófizetők jövedelemszerkezete és adótehermegoszlása 1996-ban*. TÁRKI Társadalomtudományi Tanulmányok 2.
- World Development Report* c. kiadványok, (ezen belül a „World Development Indicators” táblák). (1997), (2001) Világbank.



## SUMMARY

Using the 1998 Household Budget Survey the author compiled the incomes and expenditures of 24 socio-economic groups of the Hungarian household sector by the necessary break-down. Then, by using the recently published 1998 Input-Output Table as well as the National Accounts and the Consumption Statistics and many other statistical sources, these survey-based data were adjusted to corresponding aggregate figures of the Hungarian macrostatistics. In the case of the indirect taxes and subsidies the corresponding supplementary tables of the I-O table could be used as reference totals during the process of proportional distribution across the strata. The paper gives insight into the various methodological and technical problems of the compilation process, and presents the applied methods for reconciling the different data sources and for estimating the missing data.

Based on the compiled data base, various indicators of tax burden, and income and tax distribution are computed and presented by detailed tables. Also, the paper provides an inter-temporal and cross-country comparison of the results with the findings of previous studies.

Finally, by examples, the author illustrates, how to use the results to derive useful conclusions for economic policy making.

## A MUNKAERŐ-PIACI HELYZET A 2001. ÉVI NÉPSZÁMLÁLÁS TÜKRÉBEN

DR. FÓTI JÁNOS – DR. LAKATOS MIKLÓS

A tanulmány a munkaerő-piaci helyzet főbb jellemzőit a 2001. évi népszámlálás képviseleti mintája alapján összeállított információk felhasználásával tárgyalja. A vizsgált anyag helyes értelmezése céljából áttekintést nyújt a népszámlálás előzményeiről, a próbajellegű felvételekről, a végleges kérdőívvel kapcsolatos foglalkozási tapasztalatokról. Az elemzés kiterjed a foglalkoztatottságra, a munkanélküliségre, a foglalkoztatottak legfontosabb strukturális jellemzőire. Élve a mintavétel nagysága által biztosított lehetőségekkel, az értékelés – az országos mutatók mellett – a regionális és megyei sajátosságok taglalására is kiterjed. Tekintettel arra, hogy a 2002. évi képviseléválasztás előkészítésével kapcsolatos igények kielégítése céljából a census legfontosabb adatait az országgyűlési egyéni választókerületekre, vagyis a megyéken belüli kisebb területi egységekre vonatkozóan is feldolgozták, e kisebb területi egységek néhány speciális jellemzőit is bemutatjuk. Külön vizsgálat tárgya a Közép-magyarországi régió (Budapest, Pest megye) foglalkoztatottsági mutatóinak részletesebb területi differenciáltsága.

TÁRGYSZÓ: Népszámlálás. Foglalkoztatás. Munkanélküliség. Régió.

Az 1990-es évek közepére a gazdasági átalakulás folyamatának előrehaladása megérlelte a feltételeket ahhoz, hogy megkezdjük az előkészületeket a munkaerő-piaci viszonyoknak a foglalkoztatás szempontjából is megfelelő új szemléletű census előkészítésére. Először – 1996. április 1-jei állapot szerint – a lakosság 2 százalékára kiterjedő mintán alapuló mikrocensus (kis népszámlálás) végrehajtására került sor. Ez egyúttal a tervezett teljes körű összeírás első próbáját is jelentette, amellyel, hogy a reprezentatív adatok az adott időpontra vonatkozóan már szintén értékes információkat biztosítottak.

### A NÉPSZÁMLÁLÁST MEGELŐZŐ PRÓBAJELLEGŰ FELVÉTELEK

A mikrocensus adatai alapján a népesség gazdasági aktivitása, a foglalkoztatottság és a munkanélküliség, az aktív keresők különféle strukturális jellemzői, a munkahely és a lakóhely közötti munkaerőmozgás (ingázás) tekintetében és még számos speciális témáról készültek feldolgozások, amelyeknek eredményei rendszerint elemzést is tartalmazó kiadványok formájában nyilvánosságra is kerültek.

A minta nagysága lehetővé tette, hogy a mikrocenzus adatai országosan meglehetősen részletes információkat szolgáltatassanak, és a régiók, sőt a megyék szintjén is a fontosabb adatigényeket kellő biztonsággal kielégítsék. Nem voltak azonban alkalmasak a kistérségek és az egyes települések társadalmi–gazdasági viszonyainak a bemutatására. Ilyen mélységben ugyanis csak a teljes körű összeírás nyújthat megfelelő ismereteket.

Feltétlenül említést érdemel, hogy már a mikrocenzust megelőző időszakban is nagy jelentőségre tett szert – a megváltozott hazai követelményrendszer mellett – a nemzetközi szervezetek igényeinek teljesítése. A korábbi hazai népszámlálási gyakorlat szükségképpen a KGST igényeihez alkalmazkodott, amelyek több vonatkozásban eltértek az ENSZ, illetve a Nemzetközi Munkaügyi Szervezet (ILO) ajánlásaitól. Ismeretes például, hogy évtizedekig a munkanélküliségre vonatkozóan nem lehetett kérdést feltenni. Ebben a tekintetben már az 1990. évi népszámlálás változást hozott, mert az akkori kérdőívben 1949 óta első alkalommal szerepelt az „állást keresőkre” és az „első alkalommal állást keresőkre” vonatkozó kérdés. Az 1996. évi mikrocenzus lehetővé tette az ILO ajánlásainak megfelelő „munkanélküli” fogalom alkalmazását. Mód nyílt ezen kívül – ugyancsak a nemzetközi ajánlásoknak megfelelően – a foglalkoztatottak létszámának a kimutatására is, bár részleteiben még a régi „aktív keresők” – vagyis a nyugdíj és a gyermekgondozási ellátás mellett dolgozók létszámát nem tartalmazó – állományt mutattuk ki. (A 2001. évi népszámlálás már minden tekintetben a „foglalkoztatott” fogalmára épült.)

Az 1990-es évek elején megkezdődött az 1975 óta alkalmazott régi magyar foglalkozási osztályozás átdolgozása. E munka eredményeként a mikrocenzus az ILO rendszerének (ISCO-88) megfelelően átdolgozott korszerűsített Foglalkozások Egységes Osztályozási Rendszerét (FEOR-93) alkalmazhatta. Az ágazati nomenklatúra szintén a nemzetközi osztályozás (International Standard Industrial Classification – ISIC) 1992. évi rendszerét vette alapul. (A 2001. évi népszámlálás idején a mikrocenzushoz képest további fejlesztésre került sor. Ekkor ugyanis már rendelkezésre állt a FEOR-93 felülvizsgált, 1997-ben hatályba lépett változata, továbbá az ISIC-nek az Európai Unió tagországai számára kötelezővé tett részletesebb nomenklatúrája, a NACE (Nomenclature générale des Activités économiques dans les Communautés Européennes).

Az 1996. évi mikrocenzus eredményeinek értékelése mindenesetre nagyban hozzájárult ahhoz, hogy körvonalazódtak a népszámlálás előtt álló főbb feladatok a gazdasági aktivitás és a foglalkoztatás témakörében. Az összeírást megelőző szokásos próbafelvétel kérdéseit tehát már bizonyos mértékig konkrét ismeretekre alapozva lehetett előkészíteni. Az előkészítés során különböző változatokat dolgoztunk ki, amelyeket – az 1996. évi mikrocenzus tapasztalatainak hasznosítása mellett – szakértőkkel is egyeztettünk. Az egyeztetés egyik legfontosabb fóruma a Magyar Tudományos Akadémia Munkatudományi Bizottsága volt, amely külön ülésen tárgyalta meg a népszámlálás előkészítésével kapcsolatban felmerült kérdéseket, javaslatokat. Ezen az ülésen nemcsak a foglalkoztatottság, munkanélküliség szorosán vett tárgykörével összefüggésben vetettek fel gondolatokat, hanem más kapcsolódó témákban is, mint például a dolgozók ingázásának, vándorlásának kérdése, a szakképzettség hasznosítása, a fizikai és a szellemi fogyatékosok helyzete. Emellett a Társadalomstatistikai Tanácsadó Testület ajánlásai is lényegesen hozzájárultak a próbaszámlálás tartalmának kialakításához. Ezen kívül még számos szűkebb körű szakmai egyeztetésre, megbeszélésre is sor került a próbafelvétel teljes programjának kialakítása céljából.

Ilyen előzmények után 1999 szeptemberében két megye (Hajdú-Bihar és Heves) egyes településein mintegy 200 összeíró bevonásával, 50 ezer lakosra kiterjedően hajtottuk végre a próbanépszámlálást. Ennek programjában mind a lakáskérdőív, mind a személyi kérdőív kipróbálása szerepelt. A témánkkal kapcsolatos kérdések taglalása előtt érdemes megjegyezni, hogy a kérdőíveket – a magyar népszámlálások történetében először – a korszerű optikai karakterfelismerési rendszer (OCR) követelményeinek megfelelő formában készítették el, és a kérdőívek véglegesítésére, majd a végleges anyag feldolgozására is ebben a rendszerben került sor.

A próbaszámlálás hétoldalas személyi kérdőívének közel három oldalát a gazdasági aktivitással, foglalkozással kapcsolatos kérdések tették ki. E témakör alapját a „Megélhetését mi biztosítja?” megnevezésű kérdés jelentette, amely számos válaszlehetőséget tartalmazott. A válaszok közül legfeljebb hármat lehetett megjelölni. A válaszok módokat adtak a lakosság társadalmi–gazdasági helyzetét leginkább jellemző tényezők megállapítására, vagyis arra, hogy a számba vett népességhez tartozó személyek miként vettek részt a társadalmi munkamegosztásban.

Az elsődleges annak a kérdésnek a tisztázása volt, hogy a kérdezett személy rendelkezik-e saját jövedelemmel, és az esetleges jövedelem milyen forrásból (aktív kereső tevékenység, nyugdíjszerű ellátás, gyermeknevelési hozzájárulás, munkanélküli járadék, munkanélküliek jövedelem pótló támogatása, különböző szociális ellátások, vagyonból származó jövedelem stb.) származik. Az eltartottak esetében az eltartottság jellegére (egyéni, közületi eltartott) kellett rákérdezni. A válaszokból, illetve azok kombinációjából lényegében tisztázni lehetett, hogy a kérdezett személy foglalkoztatottnak minősül-e, vagy gazdasági aktivitás tekintetében milyen más kategóriába sorolható.

A népszámlálás hagyományos kérdései közül a foglalkozásra vonatkozó kérdés bővült. A foglalkozás, munkakör megnevezése mellett a foglalkozásra jellemző tevékenység részletesebb bejegyzésére is lehetőség nyílt. E bejegyzés a túlzottan általános vagy pontatlan foglalkozási megjelölés esetén is elősegíthette a foglalkozás helyes besorolását. Ugyancsak a pontos besorolás érdekében szerepelt a vezetői foglalkozások esetében a beosztottak számára vonatkozó kérdés.

A próbaszámlálás a szokásos heti munkaidő tudakolása mellett a részmunkaidős foglalkoztatásról is információkat kívánt szerezni. Külön kérdést tett fel arra vonatkozóan, hogy a dolgozókat szerződéssel foglalkoztatják-e, és ha igen, milyen időtartamú a szerződés.

Az összeírt személy munkáltatójáról, munkahelyéről a próbaszámlálás részletes adatokat kívánt gyűjteni. A hagyományos kérdések mellett szerepelt a munkáltató tulajdoni jellege, az esetleges külföldi érdekeltség és a munkáltatónál foglalkoztatottak létszáma. Amennyiben a munkavégzés helye valamilyen önálló részegység, részleg, telephely stb. volt, annak nevét, főtevékenységét és az ott dolgozók létszámát külön kérdés tudakolta.

A kérdőív tartalmazott kérdéseket a mellékfoglalkozásra, továbbá az adatgyűjtés időpontjában nyugdíjas vagy egyéb okból nem dolgozó személyek esetében a korábbi foglalkozásra vonatkozóan is.

A kérdések további csoportja a munkakereséssel és a munkanélküliséggel foglalkozott. E kérdések alapján meg lehetett állapítani, hogy a kérdezett személy mióta keres munkát, regisztrálták-e mint munkanélkülit, és álláslehetőség esetén mikor tud munkába állni. A témát a munkába járással kapcsolatos információk zárták. Ezek keretében tudakolták a munkahelyre jutás időtartamát, módját, a napi közlekedéshez igénybe vett járművek fajtáját.

Összességében elmondható, hogy a próbaszámlálás a foglalkozás témájában igen széles körű program végrehajtására vállalkozott. A megkérdezett lakosság válaszadási hajlandósága általában kielégítő volt. A megbízott összeírók csak kivételesen találtak elutasító magatartással. Helyenként előfordultak pontatlanságok (például a foglalkozás megnevezésénél). A munkáltató és a munkahely szétválasztása a több teleppel rendelkező nagyobb vállalatok, üzletláncok stb. esetében jelentett gondot.

### A NÉPSZÁMLÁLÁS VÉGLEGES KÉRDŐÍVÉNEK KIALAKÍTÁSA

A próbaszámlálás tapasztalatait messzemenően felhasználtuk a népszámlálás végleges programjának kidolgozásánál. Nyilvánvaló, hogy ennek során nemcsak a foglalkozási témakörre kellett tekintettel lenni, hanem a lakással, háztartással, demográfiai jellemzőkkel, iskolázottsággal stb. kapcsolatos kérdéseket is újból át kellett tekinteni. Végeredményben az a jogos észrevétel merült fel, hogy a próbaszámlálás alkalmával begyűjteni kívánt összes információ megszerzése különösen a nagyobb, összetettebb (például foglalkoztatottakat, nyugdíjasokat és tanulókat is magukban foglaló) háztartások esetében nagyon időigényes. Ennek nyilvánvalóan költségvonzatai is vannak, és emellett az összeírók és a lakosság terhelésével is számolni kell.

E szempontokat mérlegelve a Központi Statisztikai Hivatal külön munkabizottságot hozott létre, amely úgy döntött, hogy a próbaszámlálás kérdőívének a terjedelmét lényegesen – mintegy felére – kell csökkenteni. Így a foglalkozási kérdéseket is korlátozni kellett, ami a gyakorlatban egyes kérdések rövidebbé, egyszerűbbé tételét, más kérdések elhagyását jelentette. A válogatás során azonban fontos szempont volt, hogy a gazdasági aktivitás szerinti besorolás alapját képező megélhetésre, valamint munkanélküliségre vonatkozó kérdések, továbbá a dolgozók fő ágazati jellemzőinek meghatározása szempontjából nélkülözhetetlen foglalkozási és ágazati információk rendelkezésre álljanak. Nem mondhattunk le ezenkívül néhány más lényeges témáról sem, mint például a dolgozók ingázása, munkahelyre való eljutásának időtartama és módja.

A megélhetéssel kapcsolatos kérdések a három válaszlehetőség fenntartásával a kérdőív 13. és 14. pontjában szerepeltek.

13. Megélhetését mi biztosítja?					
Legfeljebb három válasz adható!					
dolgozik, jövedelmet biztosító tevékenységet végez .....	<input checked="" type="checkbox"/> 10	saját jogú öregségi nyugellátásban, járadékban részesül .....	<input checked="" type="checkbox"/> 30	munkanélküliek jövedelem-pótló támogatásában, rendszeres szociális segélyben részesül .....	<input checked="" type="checkbox"/> 51
sorkatonai vagy tartalékos katonai szolgálatot teljesít .....	<input checked="" type="checkbox"/> 11	rokkantsági nyugdíjban, baleseti járadékban részesül .....	<input checked="" type="checkbox"/> 31	egyéb segélyben, támogatásban részesül .....	<input checked="" type="checkbox"/> 60
gyermekgondozási segélyben, gyermeknevelési támogatásban részesül .....	<input checked="" type="checkbox"/> 20	hozzátartozói (özvegyi) nyugdíjban, járadékban részesül .....	<input checked="" type="checkbox"/> 40	vagyonából, egyéb forrásból származó jövedelemből él .....	<input checked="" type="checkbox"/> 70
gyermekgondozási díjban részesül .....	<input checked="" type="checkbox"/> 21	munkanélküli járadékban részesül .....	<input checked="" type="checkbox"/> 50	magánszemély (házastárs, szülő stb.) tartja el .....	<input checked="" type="checkbox"/> 80
				közületi eltartott .....	<input checked="" type="checkbox"/> 81
14. Ha magánszemély tartja el, eltartójával írták össze?					
igen, az eltartó lakáson belüli sorszáma: <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>			nem, a máshol élő eltartó fő megélhetési forrása a 13. kérdés válaszlehetőségei alapján: <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>		

A végleges kérdőívben is megmaradt a lehetőség a foglalkozás és a munkakör pontos meghatározására, a munkáltató és a munkahely szükség szerinti elválasztására, a munkáltató által foglalkoztatottak létszámának – a további feldolgozás érdekében lényeges – megállapítására, a munkahely településének megjelölésére. A kérdések e csoportja kiegészül az összeírás előtti évben végzett mezőgazdasági munkával, alkalmi, illetve idény-munkával. (Lásd a 15–18. kérdéseket.)

A 15-17. kérdéseket annak kell megválaszolnia, aki a 13. kérdésnél a 10-11. számú válaszok valamelyikét megjelölte.	
15.1 Mi a foglalkozásának, munkakörének a megnevezése, jellemző tevékenységei? ..... <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	
15.2 Van beosztottja, alkalmazottja? nincs <input checked="" type="checkbox"/> 0 van, a beosztottak, alkalmazottak száma: 1-2 fő <input checked="" type="checkbox"/> 1 3-9 fő <input checked="" type="checkbox"/> 2, 10-19 fő <input checked="" type="checkbox"/> 3, 20 fő vagy több <input checked="" type="checkbox"/> 4	
15.3 Foglalkozásában hány óra a szokásos heti munkaideje? heti <input type="checkbox"/> <input type="checkbox"/> óra kötetlen, eseti megbízás szerinti <input checked="" type="checkbox"/> 99	
16. Milyen foglalkozási formában (viszonyban) folytatja tevékenységét? alkalmazásban álló ..... <input checked="" type="checkbox"/> 1 társas vállalkozás dolgozó ..... <input checked="" type="checkbox"/> 5 segítő családtag ..... <input checked="" type="checkbox"/> 6 önálló, egyéni vállalkozó ..... <input checked="" type="checkbox"/> 2 tagja ..... <input checked="" type="checkbox"/> 3 alkalmi munkás ..... <input checked="" type="checkbox"/> 6 szövetkezet dolgozó tagja ..... <input checked="" type="checkbox"/> 4 közhasznú munkás ..... <input checked="" type="checkbox"/> 7	
17.1 Munkáltatójának, vállalkozásának mi a – pontos megnevezése? – főtevékenysége? – pontos címe? ..... <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> ..... <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	
17.2 Munkahelyének mi a – pontos megnevezése? – főtevékenysége? ..... <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	
17.3 Hol, melyik településen dolgozik? Az összeírás helyével azonos településen (kerületben) <input checked="" type="checkbox"/> 1, változó településen dolgozik <input checked="" type="checkbox"/> 2, más településen (más kerület): a település neve: ..... kerület: ..... <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	
17.4 Önnel együtt összesen hányan dolgoznak munkáltatójánál? 10 főnél kevesebb ) <input checked="" type="checkbox"/> 1, 10-19 fő ) <input checked="" type="checkbox"/> 2, 20 fő vagy több ) <input checked="" type="checkbox"/> 3, nem tudja ) <input checked="" type="checkbox"/> 4	
18.1 Az elmúlt év folyamán végzett mezőgazdasági munkát? igen, egész évben ) <input checked="" type="checkbox"/> 366, igen, az év egy részében, kb. <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> napot, nem <input checked="" type="checkbox"/> 0	
18.2 Az elmúlt év folyamán végzett alkalmi munkát, idénymunkát, dolgozott segítő családtagként? igen <input checked="" type="checkbox"/> 1, nem <input checked="" type="checkbox"/> 0	

A munkakeresésre vonatkozó kérdéskör (19. kérdés) módot nyújt mind a „hivatalos” mind az ún. passzív munkanélküliség megállapítására.

A 19. kérdésre annak kell válaszolnia, aki a 13. kérdésnél sem a 10, sem a 11 számú választ NEM jelölte meg	
19.1 Keres-e munkát? igen <input checked="" type="checkbox"/> 1, nem, mert úgy gondolja, hogy úgysem talál megfelelő munkát <input checked="" type="checkbox"/> 2, → 20. kérdés nem keres egyéb okból <input checked="" type="checkbox"/> 3 → 20. kérdés	
19.2 Mikor tudna munkába állni? két héten belül <input checked="" type="checkbox"/> 1, három-négy hét múlva <input checked="" type="checkbox"/> 2, egy-három hónap múlva <input checked="" type="checkbox"/> 3, három hónap vagy annál hosszabb idő múlva <input checked="" type="checkbox"/> 4 nem tudja <input checked="" type="checkbox"/> 5	
19.3 Mióta keres munkát? egy hónapnál rövidebb ideje <input checked="" type="checkbox"/> 0 <input type="checkbox"/> <input type="checkbox"/> hónap óta	

A társadalmi rétegződés vizsgálata szempontjából elengedhetetlennek bizonyult, hogy a népszámlálás idején már aktív kereső tevékenységet nem folytató nyugdíjasok, járadékosok stb. utolsó foglalkozását tartalmazó információk – bár a próbaszámláláshoz képest csökkentett terjedelemben – a végleges népszámlálás kérdőívén (20. kérdés) is helyet kapjanak.

A 20. kérdésre annak kell válaszolnia, aki a 13. kérdésnél a 20-81. számú válaszok valamelyikét megjelölte.	
Gyermekgondozás címén ellátásban részesülők az ellátás igénybevételét, nyugdíjasok és járadékosok a nyugdíjazást, munkanélküliek a munkanélkülivé válást megelőző foglalkozásuk adatait írják be!	
20.1 Mikor hagyta abba utolsó foglalkozását?	soha nem dolgozott <input checked="" type="checkbox"/> 0, → 21. kérdés <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> évben dolgozott utoljára
20.2 Mi volt a foglalkozásának, munkakörének megnevezése, jellemző tevékenységei?	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
20.3 Milyen foglalkozási formában (viszonyban) folytatta tevékenységét?	alkalmazásban álló <input checked="" type="checkbox"/> 1, önálló, egyéni vállalkozó <input checked="" type="checkbox"/> 2, társas vállalkozás dolgozó tagja <input checked="" type="checkbox"/> 3, szervezetet dolgozó tagja <input checked="" type="checkbox"/> 4, segítő családtag <input checked="" type="checkbox"/> 5

A terjedelem korlátozása indokolta azt a megoldást is, hogy a munkahelyre, valamint az iskolába közlekedők egy kérdéscsoportba (21. kérdés) kerültek oly módon, hogy mind a munkahely, mind az iskola esetében legfeljebb három közlekedési eszközt lehetett megjelölni.

21.1 Kell közlekednie napi rendszerességgel?		iskolába: nem <input checked="" type="checkbox"/> 0, igen <input checked="" type="checkbox"/> 1			
munkahelyre: nem <input checked="" type="checkbox"/> 0, igen <input checked="" type="checkbox"/> 1					
21.2 Általában hogyan jut el Legfeljebb három-három közlekedési eszközt jelölhet meg.					
munkahelyére		iskolába			
gyalog	<input checked="" type="checkbox"/> 1	<input checked="" type="checkbox"/> 1	távolsági autóbusz	<input checked="" type="checkbox"/> 6	<input checked="" type="checkbox"/> 6
villamoson, trolibuszon	<input checked="" type="checkbox"/> 2	<input checked="" type="checkbox"/> 2	vonaton	<input checked="" type="checkbox"/> 7	<input checked="" type="checkbox"/> 7
metrón, földalatin	<input checked="" type="checkbox"/> 3	<input checked="" type="checkbox"/> 3	autón	<input checked="" type="checkbox"/> 8	<input checked="" type="checkbox"/> 8
gyorsvasúton, hév-en	<input checked="" type="checkbox"/> 4	<input checked="" type="checkbox"/> 4	kerékpáron, motorkerékpáron	<input checked="" type="checkbox"/> 9	<input checked="" type="checkbox"/> 9
helyi autóbusz	<input checked="" type="checkbox"/> 5	<input checked="" type="checkbox"/> 5	egyéb módon	<input checked="" type="checkbox"/> 10	<input checked="" type="checkbox"/> 10
21.3 Általában naponta hány percet vesz igénybe a közlekedés a munkahelyre, iskolába (a visszaúttal és a várakozással együtt)?					
munkahelyre: <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> percet		iskolába: <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> percet			

## A NÉPSZÁMLÁLÁS VÉGREHAJTÁSÁNAK TAPASZTALATAI

A csökkentett terjedelmű személyi kérdőív a 2001. február 1-jei állapot szerint végrehajtott hivatalos népszámlálás során felhasználásra került. Az összeírást végző számlálóbiztosok, valamint a munkájuk ellenőrzésére kijelölt felülvizsgálók a munka megkezdése előtt országsszerte oktatáson vettek részt, amelynek keretében megismerkedhettek a kérdőívvel és a kitöltés szabályaival. Tennivalóikat írásban kiadott részletes utasítás is tartalmazta. A foglalkozási témakörnél – a népszámlálás más témaköreire is hasonlóan – számos sajátos probléma adódott. Ezekre külön példákat dolgoztunk ki, amelyek felhasználása az oktatás során jelentős mértékben elősegítette az összeíráskor jelentkező gyakorlati feladatok megoldását.

Meg kell ugyanakkor jegyezni, hogy az oktatásra fordítható idő rövidsége miatt nem mindenütt nyílt lehetőség a problémák kellő mélységű megvitatására. Felmerülhettek ezen kívül olyan sajátos helyi gondok, amelyeket az oktatás, illetve a felkészülés időszakában még nem lehetett ismerni. Érthető ezért, hogy az összeírás időszakában is felmerültek kérdések, amelyeket szakértői állásfoglalások, pontosítások kiadásával kellett rendezni. (A korábbi censzusok és általában a lakossági összeírások tapasztalatai azt mutatják, hogy még a legalaposabb oktatás mellett is sor kerülhet sajátos helyi problémák felvetésére, amelyek csak az összeírás folyamatában, „menet közben” oldhatók meg.) Végeredményben megállapítható, hogy a foglalkozási témakörbe tartozó kérdések megválaszolása az esetek döntő többségében megfelelő volt. Helyenként adódtak hibák és hiá-

nyosságok, amelyek többnyire egyes számlálóbiztosok felületességéből, valamint egyes kérdések téves értelmezéséből adódtak. A lakosság válaszadási készsége általában kielégítőnek bizonyult. Bizalmatlanságot csak elvétve lehetett tapasztalni. A főváros néhány kerületében, valamint egyes nagyobb településeken fordult inkább elő, hogy némely kérdésre a felkeresett személyek nem szívesen válaszoltak, illetve hiányos választ adtak.

Az utasítás értelmében a 13. kérdésre (Megélhetését mi biztosítja?) mindenképpen válaszolni kellett, vagyis a feltüntetett válaszlehetőségek közül legalább egyet meg kellett jelölni. Ennek ellenére helyenként – bár viszonylag ritka esetekben – előfordult, hogy a válaszadás elmaradt. Amennyiben viszont a foglalkozásra, a munkáltatóra stb. vonatkozó válaszokat kitöltötték, a feldolgozás során 15 éven felüli személynél az aktív kereső tevékenységre utaló 10. választ, 15 éven aluliaknál az eltartottságra utaló válaszlehetőségeket (80, 81) vélelmezni lehetett.

A foglalkozást, munkakört tudakoló 15.1 kérdésnél egyes számlálóbiztosok anyagának ellenőrzése során szembesülnünk kellett azzal a problémával, hogy a foglalkozási besorolásra alkalmatlan, általános jellegű bejegyzések (például osztályvezető, tisztviselő, betanított munkás) szerepeltek. Az ilyen jellegű hibák elkerülése érdekében – a próbaszámlálás tapasztalatait is mérlegelve – az oktatás alkalmával külön felhívtuk a figyelmet a pontos foglalkozási megnevezések használatára, és példákkal is bemutattuk a foglalkozások helyes meghatározásának módját. Mindez nyilvánvalóan csökkentette az ilyen hibák előfordulását, de azokat teljesen nem küszöbölhette ki. Számolni kell azzal is, hogy önkitöltésre is lehetőség volt, és az önkitöltők nem fordítottak mindig kellő gondot a foglalkozás pontos bejegyzésére.

A fő gondot a 17.1 kérdés (Munkáltatójának, vállalkozásának, mi a – pontos megnevezése? – fő tevékenysége? – pontos címe?), továbbá a 17.2 kérdés (Munkahelyének mi a – pontos megnevezése? – fő tevékenysége?) megválaszolása okozta.

Ismeretes, hogy a 2001. évi népszámlálás kérdőívének kitöltése eleve név nélkül történt. Az összeírás kezdetén – az azonosíthatósággal kapcsolatos aggodalmak kiküszöbölése érdekében – olyan intézkedés született, hogy a lakcímet sem kell feltüntetni. A foglalkoztatottak esetében a munkáltató (vállalkozás) pontos megnevezése, címe azonban elvben továbbra is lehetőséget adhatott az azonosításra. Helyenként ezért tapasztalni lehetett azt a feltételezést, hogy ezek az információk az adóhatóságok vagy más illetéktelen szervezetek, személyek kezébe kerülhetnek. Az összeírók figyelmét ismételtelen felhívtuk arra, hogy a válaszadást a jelzett ok miatt megtagadókat hangsúlyozottan tájékoztassák a Központi Statisztikai Hivatalnak az adatok bizalmas kezelésével kapcsolatos kötelezettségéről, az adatvédelmi szabályokról, amelyek tiltják egyedi adatok kiszolgáltatását. A bizalmatlanságot mégsem sikerült mindenütt teljesen kiküszöbölni. Egyes munkáltatók kifejezetten megtiltották alkalmazottaiknak az adatszolgáltatást. Sajnálatos módon a média egy részének magatartása is fokozta a negatív hozzáállást. Szerencsére az ágazati besorolás alapját képező fő tevékenység bevallása kevésbé ütközött ellenállásba. A teljes körű összeírás alkalmával is néhol tisztázni kellett a munkahely meghatározásának, a munkáltató és munkahely kapcsolatának kérdését, noha – a próbaszámlálás tapasztalatait hasznosítva – az oktatás keretében számos példával bemutattuk a tipikusnak tekinthető esetek megoldását.

A munkakeresésre vonatkozó 19. kérdéscsoportot illetően a válaszadási készség általában megfelelő volt. A válaszok alapján meg lehetett állapítani – összhangban a nemzetközi



ajánlásokkal – a ténylegesen munkanélküli személyek, továbbá az ún. „passzív munkanélküliek” (az állást keresők, de az állás megszerzését reménytelennek ítélik) számát.

A korábban dolgozó, de jelenleg nyugdíjból, járadékból vagy egyéb forrásból élő személyek utolsó foglalkozásának bejegyzésénél (20. kérdéscsoport) még határozottabban mutatkoztak azok a pontatlanságok, amelyeket a jelenleg dolgozók foglalkozásával kapcsolatban említettünk. Itt természetesen jelenleg már nem létező foglalkozások, munkakörök (például tanácselnök, tanácsítók, üzemi párttitkár) feltüntetésével is számolni kellett.

A napi közlekedési szokásokat tudakoló 21. kérdéscsoportnál a leggyakoribb hiba a kétféle (munkahelyre, illetve iskolába irányuló) közlekedés felcserélése volt.

### AZ ÖSSZEÍRÁS ANYAGÁNAK FELDOLGOZÁSA ÉS PUBLIKÁLÁSA

A feldolgozás első munkafázisa a kódolás volt. A 2001. évi népszámlálás a korábbiaktól eltérő gyakorlatot követett: a kódolás egy részét a felülvizsgálónak még az anyag népszámlálási felelősnek való átadása előtt kellett elvégeznie. Ez az ún. előkódolás a foglalkozási kérdéskört csak a 17.3 kérdés vonatkozásában érintette. Amennyiben a munkahely az összeírás helyével nem azonos településen (Budapesten nem azonos kerületben) volt „A községek, városok és az országok betűrendes jegyzéke” alapján a település (kerület) nevének megfelelő ötjegyű számot kellett bejegyezni. A fővárosi munkahelyek esetében gyakori hiba volt, hogy a kerület kódolása akkor is elmaradt, amikor a munkahely más kerületben szerepelt, mint az összeírás helye.

A kódolás további döntő és részben bonyolult lépéseire már a leadott anyagon került sor. A kódolást részleges kódrevízió és szuperrevízió követte. A végleges javításra az OCR-feldolgozás keretében került sor. A foglalkozási és munkáltatói adatok kódolása elkerülhetetlenül ellenőrzési és vélelmezési feladatokkal is járt. Ehhez a rendszeres szakmányfelvilágosítás nyújtott segítséget; ezen kívül a gyakrabban felmerült kérdésekről külön kiegészítő anyagokat is kiadtunk.

A 15.1 kérdésnél bejegyzett foglalkozási megnevezések sokszor helyi sajátosságokat tükröztek, és nem mindig feleltek meg pontosan „A foglalkozások jegyzéke” c. segédletben szereplő megnevezéseknek, amelyek a FEOR-93 rendszerére épültek. Rendszerint azonban logikai úton mód nyílt az adott foglalkozással, munkakörrel tartalmilag azonosítható kódszám meghatározására. Nagyobb gondot jelentettek a már említett túlzottan általános, tényleges információt nem hordozó bejegyzések. Ilyen esetekben a vélelmezést az adott munkahely fő tevékenysége, esetleg a dolgozó iskolázottsága, szakképzettsége segíthette.

A vezetői munkakörök kódolásánál viszonylag gyakran mutatkozott bizonytalanság. A kódolók egy része kritika nélkül elfogadta a „vezető” bejegyzést, ahelyett hogy az utasítás értelmében megvizsgálta volna, hogy ténylegesen vezetői funkcióról van-e szó, és tekintetbe vette volna – mint lényeges ismérvet – a beosztottak számát. Számos esetben került sor például az 1311 (ügyvezető igazgató) kódszám irreális alkalmazására. Ellenkező irányú tévedések is előfordultak, amikor vezető beosztású dolgozót nem vezetőként kódoltak.

A foglalkozás kódolásánál még több kisebb-nagyobb probléma jelentkezett. A teljesség igénye nélkül meg kell említeni például, hogy egyes kódolók nem a megfelelő segédletet használták: például a foglalkozási kódszámokat a foglalkozási jegyzék helyett az

iskolázottságra, szakképzettségre vonatkozó jegyzékből vették. E tévedés szerencsére nem volt gyakori, de zavart okozott, mert hibás kódszámokat eredményezett. A kódolás első szakaszát követően a kódolók az ilyen természetű hibákról rendszerint visszajelzést kaptak, úgy hogy a későbbiekben ezek előfordulása lényegesen csökkent.

A legtöbb gondot kétségkívül a munkáltató és a munkahely kódolása jelentette. A korábbi népszámlálások alkalmával a munkáltatók jegyzékének összeállítása a vállalatok, intézmények viszonylag kis száma miatt nem okozott érdemi nehézséget. A rendszerváltozás eredményeként a gazdálkodó szervezetek száma sokszorosára emelkedett. Ezekről teljes körű jegyzék összeállítása szóba sem jöhetett. Végül azt a megoldást választottuk, hogy a munkáltatók egyedi címeit tartalmazó jegyzék csak a 20 főt vagy annál nagyobb létszámot foglalkoztató szervezetekről készült. (A jegyzék így is közel 20 ezer címet tartalmazott.) E szervezetek esetében e munkáltatói jegyzékből kellett a kódszámot megállapítani. Az adatok pontosabb meghatározása érdekében az ide sorolt szervezetek egyedi azonosítására is lehetőséget biztosítottunk. Ennek érdekében a kódolónak – ha az adott munkáltatót a jegyzékben megtalálta – a név mellé írt öt számjegyű azonosítót is be kellett jegyeznie a megfelelő kódnégyzetekbe. Amennyiben a foglalkoztatottak száma nem érte el a 20 főt, „A gazdasági tevékenységek jegyzéke” című segédlet képezte a kódolás alapját. Az utóbbi eljárást kellett követni akkor is, ha a legalább 20 fősnek jelzett munkáltató nem szerepelt a munkáltatói jegyzékben. Az a körülmény, hogy kétféle jegyzéket kellett használni, nehezítette a kódolást, különös tekintettel arra, hogy minden egyes esetben a munkáltató létszámát is figyelni kellett, mert csak annak alapján lehetett eldönteni, hogy melyik jegyzék alkalmazható.

A munkáltatói jegyzék használatát kétségkívül korlátozta az a körülmény, hogy az a gazdálkodó szervek regisztere alapján készült; ezért a regiszter esetleges hibáit, hiányait is átvette. Előfordulhatott az is, hogy az adott cég hivatalos nevén szerepelt a jegyzékben, de az a bemondott fantázianévet, rövidítést nem tartalmazta. A kódoló tehát a ténylegesen létező céget hiába kereste a munkáltatói jegyzékben. A kódolók nagy része jelentős energiát fordított arra, hogy a kérdőíven szereplő bejegyzés alapján megtalálja a munkáltatót. Egyesek azonban az időigényesség miatt a könnyebb megoldást választották: egyszerűen a tevékenységi jegyzék alapján végezték a kódolást. Amennyiben a fő tevékenység bejegyzése pontos volt, ez az ágazati besorolás helyességét elvben nem befolyásolta. A munkáltatói jegyzékben feltüntetett ötjegyű azonosító bejegyzése azonban ilyen esetekben elmaradt ott is, ahol megtalálható lett volna.

Az ismertetett példák is jelzik, hogy a foglalkozási és a munkáltatói adatok kódolása során számos probléma merült fel, amelyeket a revízió is csak részben tudott megoldani. Végül az OCR-feldolgozás – amelyben foglalkozásstatisztikai szakértők is részt vettek – döntő részben pótolta a még fennmaradt hiányosságokat, és lényeges mértékben javítani tudta az adatok minőségét.

#### A 2001. ÉVI NÉPSZÁMLÁLÁS ADATAINAK PUBLIKÁLÁSA

A népszámlálási adatok feldolgozásának folyamata még a kezdeti szakaszban tartott, amikor – a korábban követett gyakorlatnak megfelelően – megjelent az előzetes adatokat tartalmazó első kiadvány. E kiadvány a számlálókörzeti gyűjtőívek adatainak alapján magában foglalta a lakásokra, valamint a lakók számára vonatkozó információkat teljes

területi részletezettségben, utalva arra, hogy az adatok a végleges feldolgozás során területenként kisebb-nagyobb mértékben módosulhatnak.

Az előzetes összesítésben még nem szerepelhettek sem a gazdasági aktivitásra, foglalkozásra, sem a részletes demográfiai, iskolázottsági jellemzőkre stb. vonatkozó információk. A fontosabb ismérvek mielőbbi közzététele érdekében, mintegy 1,3 millió főre kiterjedő reprezentatív feldolgozás készült.<sup>1</sup> Ennek főbb eredményeit a még 2001 decemberében napvilágot látott második, „Részletes adatok” című kötet tartalmazza, az országos összesítésen kívül regionális és megyei bontásban.<sup>2</sup> Az adatok publikálásának megtervezésénél tekintetbe kellett venni, hogy a 2002. évi választások alkalmával a megyéken belül az országgyűlési egyéni választókerületekre vonatkozó alapvető információk is érdeklődésre tarthatnak számot. A minta nagyságát, felépítését ezért úgy alakítottuk, hogy a leglényegesebb arányok még e kisebb térségek helyzetét is kellő megbízhatósági szinten reprezentálják. Ezen igények kielégítésére „Az országgyűlési egyéni választókerületek adatai” címmel 20 füzet készült.<sup>3</sup> E füzetek a képviselői mintán alapuló adatokat megyénként és külön a fővárosra tartalmazzák. Hangsúlyozni kell, hogy a választókerületek nem azonosíthatók a statisztikai kistérségekkel, hiszen egészen más célt szolgálnak. Az országban 176 egyéni választókerület és 150 statisztikai kistérség van. Az egyéni választókerületek és a kistérségek átlagos mérete azonban nem tér el érdemlegesen. Egy-egy választókerület adatai ezért megközelítőleg jellemezhetik a megyén belüli térség helyzetét, és – természetesen kellő óvatossággal – bizonyos kistérségek sajátos jellemzőire vonatkozó következtetések levonását is lehetővé teszik.

A továbbiakban a foglalkoztatottsági jellemzőket először országosan, majd területi szinten tesszük vizsgálat tárgyává, az említett minta feldolgozásából rendelkezésre álló információk alapján.

### *Országos adatok*

A népesség *foglalkoztatottságának, gazdasági aktivitásának* változása lényegében már az 1980-as években megindult, a XX. század utolsó évtizedében, vagyis a rendszerváltozás, a társadalmi és gazdasági átalakulás időszakában pedig alapvető strukturális átrendeződésre került sor. Az adatok értékelésénél utalnunk kell arra, hogy a gazdasági aktivitást jellemző főbb folyamatok erősen kötődtek a rendszerváltozás menetéhez. Az 1990-es évtized első felében a foglalkoztatottak száma igen nagy mértékben, kereken egymillió fővel csökkent. Ez a gyors ütemű privatizációnak, a létszámleépítéseknek, a gazdaságtalan vállalkozások megszűnésének volt a következménye. Az évtized közepétől a visszaesést stagnálás, majd a foglalkoztatottak számának lassú növekedése váltotta fel. Ennek ellenére 2001. február 1-jén még mindig közel egyötöddel kisebb volt a foglalkoztatottak száma, mint tizenegy évvel korábban.

A foglalkoztatottság csökkenésének egyik fő következménye az 1990-es évtized első éveiben a munkanélküliek számának rohamos növekedése volt, ami – a regisztrált munkanélküliek adata szerint – 1993-ban érte el csúcspontját. A regisztrált munkanélküliek létszáma ezt követően némileg csökkent, de még az évtized közepén is magas volt: az

<sup>1</sup> Ennek módszerét lásd bővebben: Népszámlálás 2001. 2. Részletes adatok a képviselői minta alapján. (559–565. old.)

<sup>2</sup> Lásd interneten: [www.ksh.hu](http://www.ksh.hu) vagy [www.nepszamlalas2001.hu](http://www.nepszamlalas2001.hu).

<sup>3</sup> Lásd interneten: [www.ksh.hu](http://www.ksh.hu) vagy [www.nepszamlalas2001.hu](http://www.nepszamlalas2001.hu).

1990. évinek mintegy négyszeresét tette ki és a 2001. évi népszámlálás is az 1990. évi munkanélküli-állománynak több mint háromszorosát rögzítette. A munkanélküliségi ráta az 1990. évi 2,7 százalékkal szemben 10,2 százalék volt.

A 2001. évi népszámlálás munkanélküliségi adatai magasabbak, mint a KSH folyamatos munkaerő-felmérésének erre vonatkozó adatai. Ennek egyik oka, hogy a népszámlálás – az ILO fogalmi rendszerén belül maradván ugyan – a kérdőív terjedelmi és a kikérdezési idő korlátai miatt nem vizsgálhatta olyan részletesen a munkanélküliséget, mint a munkaerő-felmérés. Ez utóbbinál ugyanis mód van arra, hogy a munkanélküliség mértékét több kérdés feltevésével számos más kritériumot is figyelembe véve állapítsák meg. A munkaerő-felmérésben például azt a személyt, aki a felmentési idejét tölti, a foglalkoztatottak közé számítják akkor is, ha munkanélkülinek mondja magát. A népszámlálásnál ilyen korrekcióra nem volt lehetőség, mert a kérdőíven nem szerepelt a felmondásra vonatkozó kiegészítő kérdés. A népszámlálás tehát inkább a kérdezett személyek szubjektív megítélésén alapuló állapotot rögzítette. A népszámlálás munkanélküliségi adatai azonban a számszerű eltérés ellenére strukturális szempontból hasonló tendenciákat mutatnak, mint a munkaerő-felmérésből származó információk. Az adatok így egyfelől a munkanélküliség területi különbségeit, a munkanélküliek demográfiai, háztartási, családi és lakásjellemzőit is bemutatják, másfelől biztosítják a más országok népszámlálási adataival való összehasonlítást.

Az előző évtized gazdasági és társadalompolitikai gyakorlatának másik lényeges következménye, hogy megnőtt az inaktív keresők száma. A rendszerváltozást követő években meglehetősen széles körben tették lehetővé az akkori nyugdíjkorhatárt (férfiaknál 60 év, nőknél 55 év) még el nem ért, de azt megközelítő dolgozóknak a korengedményes nyugdíj, majd az előnyugdíj igénybevételét. Emellett a piacgazdaság térhódítása következtében csökkent a valamilyen egészségkárosodásban szenvedő munkavállalók szervezett foglalkoztatásának a lehetősége, e munkavállalók közül sokan rokkantsági nyugdíjasok lettek. Mások – akiket egészségügyi problémáik a „teljes foglalkoztatottság” időszakában ténylegesen nem gátoltak az aktív kereső tevékenységben – a létszámleépítések hatására a munkanélküliség helyett inkább a rokkantsági nyugdíjazás lehetőségét választották.

Végeredményben – a fiatalabb korcsoportokat is érintően – erőteljesen bővült a rokkantsági nyugdíjban részesülők köre. A nyugdíjban, illetve nyugdíjrendszerű ellátásban részesülők számának emelkedése az utóbbi néhány évben mérséklődött, sőt az öregségi nyugdíjasok csoportjában legújabbban már számszerű csökkenést idéz elő a nyugdíjkorhatár fokozatos emelkedése (2001-ben a férfiaknál 61 év, a nőknél 57 év lett a korhatár).

Az inaktív keresők sajátos rétegét alkotják a gyermeknevelési segélyezésben részesülők. A körükben tapasztalható átmeneti létszámcsökkenést a gyermekgondozási díj visszaállítása megszüntette.

Az eltartottak köre 1990 óta fokozatosan kisebb lett, aminek alapvető oka, hogy a születésszám évtizedek óta tapasztalt kedvezőtlen alakulása következtében alacsonyabb a munkavállalási koron aluli, valamint a döntően nappali tagozatos tanulmányokat folytató 15–19 éves korúak száma.

Az ismertetett tényezők a népesség gazdasági aktivitás szerinti összetételét az 1. táblában bemutatott mértékben módosították.

1. tábla

*A népesség megoszlása 1990-ben és 2001-ben gazdasági aktivitás szerint*

Év	Összes	Foglalkoztatottak	Munkanélküliek	Inaktív keresők	Eltartottak
1990	100,0	43,6*	1,2	25,7	29,5
2001	100,0	36,2	4,1	33,3	26,4

\* Itt és a továbbiakban a nyugdíj és a gyermekgondozási ellátás mellett dolgozók száma nélkül számított arány.

Az adatok jól mutatják a két népszámlálás között, tizenegy év alatt végbement jelentős átrendeződést. 2001-ben több mint 7 százalékponttal volt kisebb a foglalkoztatottak aránya, mint 1990-ben. Ennek következtében lényegesen megnőtt a foglalkoztatottak terhelése, mivel az 1990. évi 129-cel szemben 2001-ben már 176 aktív tevékenységet nem folytató személy jutott száz foglalkoztatottra. A foglalkoztatottak mellett a nem dolgozó, de munkavégzésre rendelkezésre álló munkanélkülieket is magában foglaló gazdaságilag aktív népesség aránya összességében 4,4 százalékponttal csökkent. Ezzel szemben áll az inaktív keresők népességén belüli részesedésének közel 8 százalékpontos növekedése.

A teljes foglalkoztatás elvének feladása után sokan arra számítottak, hogy az aktív kereső nők tömegesen adják fel kereső foglalkozásukat, és térnek vissza a háztartásba. A nők visszavonulása az aktív kereső tevékenységtől azonban nem öltött a feltételezéseknek megfelelő tömeges méreteket. Gazdasági aktivitásuk a férfiakéval közel azonos ütemben esett vissza. Ebből adódott, hogy a nők aránya a foglalkoztatottak körén belül 2001-ben 45,6 százalékot tett ki, közel egy százalékponttal meg is haladva az 1990. évi szintet (44,5%).

A foglalkoztatott nők számában bekövetkezett viszonylag mérsékelt csökkenés főként azzal magyarázható, hogy a munkanélküliség a nőket kevésbé érintette, mint a férfiakat. E jelenség már 1990-ben is megfigyelhető volt, 2001-ben pedig fokozottan érzékelhető a különbség: a férfiaknak 5,2, a nőknek 3,1 százaléka munkanélküli.

2. tábla

*A népesség gazdasági aktivitása korcsoportok szerint 2001-ben*

Korcsoport	Összes	Foglalkoztatottak	Munkanélküliek	Inaktív keresők	Eltartottak
0-14	100,0	0,0	0,0	0,0	100,0
15-19	100,0	8,5	4,1	2,6	84,8
20-29	100,0	58,6	8,8	16,3	16,3
30-39	100,0	70,5	7,9	18,0	3,6
40-49	100,0	71,6	6,6	18,5	3,3
50-59	100,0	47,5	3,3	46,9	2,3
60-69	100,0	4,7	0,3	94,0	1,0
70-74	100,0	1,3	0,0	97,3	1,4
75-	100,0	0,8	0,0	97,8	1,4
<i>Összesen</i>	<i>100,0</i>	<i>36,2</i>	<i>4,1</i>	<i>33,3</i>	<i>26,4</i>

A 2001. évi népszámlálás gazdasági aktivitással kapcsolatos adatai a 15 éves és idősebb népességre vonatkoznak. (A 15 éven aluli népesség kizárólag eltartottakból tevődik össze.) A 2. tábla korcsoportonkénti adatai arra hívják fel a figyelmet, hogy a foglalkoztatottság azokban a korcsoportokban, azaz a 20–29 éves fiatalok, valamint a középgenerációba tartozó 30–39 évesek és 40–49 évesek csoportjában is viszonylag alacsony, amelyekre hagyományosan a foglalkoztatottak magas aránya volt jellemző.

A 20–29 éveseknek kerekén 59 százaléka foglalkoztatott. Ez 1990-hez képest 16 százalékponttal csökkent, amiben kétségkívül szerepet játszik, hogy jelentős mértékben megemelkedett a nappali tagozaton egyetemi, főiskolai tanulmányokat folytatók száma. Úgyszintén erre a korcsoportra jellemző leginkább, hogy gyermeknevelés miatt sokan átmenetileg nem folytatnak aktív kereső tevékenységet. Figyelmet érdemel, hogy a 20–29 évesek között – az átlagos szint kétszeresét is meghaladóan – a legmagasabb a munkanélküliek aránya.

A középgenerációban a foglalkoztatottság 70-72 százalék körüli szintje szintén mérsekelt értéknek tekinthető. Az ide sorolt két korcsoportba tartozóknak közel egyötöde gazdaságilag inaktív, amiben már lényegesen érzékelhető a munkanélküliség helyett a rokkantsági nyugdíjba menekülés.

Az 50–59 éves korcsoportba tartozóknak már csak kevesebb mint fele foglalkoztatott, ugyanakkor az inaktív keresők hányada majdnem akkora, mint a foglalkoztatottaké. Tekintettel arra, hogy a jelenlegi szabályozás szerint az 50–59 éves férfiak teljes köre és az ilyen korú nőknek többsége még nem jogosult öregségi nyugdíjra, e magas arány kialakulásában lényeges szerepük van a különböző kedvezményeknek (például korengedményes nyugdíj, előnyugdíj).

A dolgozók *foglalkozási főcsoport* szerinti összetétele huzamosabb idő óta folyamatos átalakulást mutat.

3. tábla

*A foglalkoztatottak megoszlása összevont foglalkozási főcsoportok szerint*

Összevont foglalkozási főcsoport	1990.	2001.
	évben (százalék)	
Vezető, értelmiségi foglalkozású	16,1	20,7
Egyéb szellemi foglalkozású	17,1	20,0
Szolgáltatási foglalkozású	8,6	15,2
Mezőgazdasági foglalkozású	4,0	3,2
Ipari, építőipari foglalkozású	40,2	31,3
Egyéb foglalkozású	14,0	9,6
<i>Összesen</i>	<i>100,0</i>	<i>100,0</i>

Az elmúlt évtizedben ez főleg abban nyilvánult meg, hogy a foglalkoztatottak között a vezető, értelmiségi és az egyéb szellemi foglalkozásúak aránya fokozatosan emelkedett. 1990-ben a foglalkoztatottaknak egyharmada, 2001-ben 41 százaléka folytatott szellemi tevékenységet. (Lásd a 3. táblát.) E tendencia azonban nem csupán a foglalkoztatottak iskolázottsági szintjének általános emelkedését mutatja, hanem azt is, hogy az elmúlt évti-

zedben a foglalkoztatottak számának csökkenése – elsősorban a munkanélkülivé válás következtében – főleg a fizikai foglalkozásúakat érintette. A fizikai foglalkozásúak számának csökkenése egyben azt is jelzi, hogy az ország gazdasági szerkezete egyre inkább a képzetlenebb munkaerő alkalmazását igénylő gazdasági ágazatok, tevékenységek irányába mozdult el. A szolgáltatási jellegű tevékenységet folytató foglalkoztatottak arányának jelentős emelkedése, a mezőgazdaság és erdőgazdálkodás, valamint az ipar és építőipar területén dolgozók arányának visszaesése a foglalkozás oldaláról is bizonyítható.

A mezőgazdasági, illetve az anyagi termeléshez közvetlenül kapcsolódó munkafolyamatokban a férfiaknak továbbra is meghatározó szerepük van. A segédmunkások döntő többségét magukba foglaló egyéb foglalkozásúak körében a nemek aránya szintén kiegyenlítettnek mondható, e körben vannak olyan foglalkozások (például rakodómunkás), amelyeket nagyrészt férfiak, és vannak olyan tevékenységek (például takarítás, konyhai kisegítés), amelyeket továbbra is inkább nők végeznek.

Összevont foglalkozási főcsoportok és korcsoportok szerint a foglalkoztatottak aránya jelentős korszpecifikus eltéréseket mutat. A fiatalok (15–29 évesek) aránya kiugróan magas a szolgáltatási foglalkozásúak körében, alacsony viszont – a szakmai gyakorlat hiányából adódóan – a vezető és értelmiségi foglalkozásúak – ezen belül különösen a vezetői munkakört betöltők – között. A mezőgazdasági foglalkozásúak átlagosnál magasabb korösszetétele ebben a csoportosításban is kimutatható. A magasabb szellemi képzettséget igénylő vezetői, értelmiségi foglalkozásúak között viszonylag magas a nyugdíjkorhatárt már betöltött foglalkoztatottak aránya.

A foglalkoztatottak, foglalkozás szerinti megoszlása az iskolai végzettséggel is összefügg. Érthető, hogy a magas képzettséget igénylő vezető, értelmiségi foglalkozásúak körében legmagasabb (69%) a felsőfokú iskola legalább első évfolyamát végzettek aránya. Mivel a gyakorlatban (mindenekelőtt a versenyszférában) a vezetővé válás nem feltétlenül függ a megszerzett bizonyítványtól, oklevéltől, a vezetők esetében nem alapvető szempont az iskolázottság, ennek mértéke elsősorban az értelmiségi foglalkozásoknál magas. A középiskola 9–15. osztályát végzettek aránya nagyon magas az egyéb szellemi foglalkozásúak (78%) és az ipari, építőipari tevékenységet végzők (79%) körében. Az előbbieken esetében ez inkább gimnáziumi, szakközépiskolai, az utóbbiaknál főleg szakközépiskolai, szakmunkásképző-iskolai végzettséget jelent. Az egyéb foglalkozásúak nagy részét kitevő segédmunkás-jellegű foglalkozásúak döntő többsége (73%) csak az általános iskola 8. vagy annál alacsonyabb osztályát végezte el.

Ha egy-egy összevont nemzetgazdasági ágat (ágakat) abból a szempontból vizsgálunk, hogy az ott dolgozó foglalkoztatottak milyen jellegű tevékenységet folytatnak, akkor a legegységesebb nemzetgazdasági ág az ipar és építőipar, ugyanis ebben az összevont nemzetgazdasági ágban dolgozó ipari, építőipari foglalkozásúak aránya 65 százalék. A mezőgazdaság és erdőgazdálkodás dolgozóinak körében a mezőgazdasági foglalkozásúak aránya már csak 50 százalék. Az egymástól leginkább eltérő jellegű foglalkozások a szolgáltatási ágakban találhatók.

A foglalkoztatottak *nemzetgazdasági ág szerinti összetétele* már az 1980-as évtized során is jelentős mértékben módosult. A korábban is tapasztalt tendenciákat pedig az 1990 óta végbement társadalmi és gazdasági átalakulás felgyorsította.

A mezőgazdasági szektor válsága és az ipari termelés visszaesése gyökeresen megváltoztatta a foglalkoztatottak ágazati megoszlását: 2001-re a mezőgazdaság és erdőgaz-

dálkodás (vadgazdálkodással és halászattal együtt) részesedése 1990-hez képest 10 százalékponttal (15,5-ről 5,6-re) esett vissza, az ipar (bányászat, feldolgozóipar, villamosenergia-ipar, gáz-, gőz- és vízellátás) és építőipar részesedése közel 5 százalékponttal, alig egyharmadra (37,9-ről 33,1-re) csökkent, ugyanakkor a szolgáltatás jellegű ágazatok (az előbbieken nem említett nemzetgazdasági ágak együtt) súlya jelentősen (46,7-ről 61,3-re) nőtt. Amíg 1990-ben a foglalkoztatottak 47 százaléka, addig 2001-ben már több mint 61 százaléka dolgozott a tercier ágazatokban. Az arányváltás önmagában gyors modernizációs folyamatot tükröz, amely strukturális átalakulás a foglalkoztatottak számának igen jelentős csökkenése mellett ment végbe.

Az összevont nemzetgazdasági ágakban foglalkoztatottak nemek szerinti összetétele jelentős eltéréseket mutat. A férfiak foglalkoztatottakon belüli aránya változatlanul magasabb azokban a nemzetgazdasági ágakban – különösen a mezőgazdaságban és erdőgazdálkodásban –, ahol még mindig jelentős szerepe van a fizikai erőfelfejtésnek. A szolgáltatás jellegű ágazatokban viszont a nők hányada a magasabb.

4. tábla

*A foglalkoztatottak százalékos megoszlása összevont nemzetgazdasági ágak szerint 2001-ben*

Összevont nemzetgazdasági ágak	Összesen	Férfi	Nő	15–29	30–39	40–49	50–59	60–
				éves				
Mezőgazdaság és erdőgazdálkodás	100,0	75,2	24,8	17,9	23,1	34,0	22,3	2,7
Ipar és építőipar	100,0	65,6	34,4	28,4	24,4	30,1	16,3	0,8
Szolgáltatás jellegű ágazatok	100,0	46,3	53,7	26,5	25,7	29,3	16,6	1,9
<i>Összesen</i>	<i>100,0</i>	<i>54,4</i>	<i>45,6</i>	<i>26,6</i>	<i>25,1</i>	<i>29,9</i>	<i>16,8</i>	<i>1,6</i>

A foglalkoztatottak többségét mindegyik összevont nemzetgazdasági ágban a középkorúak (30–49 évesek) teszik ki. (Lásd a 4. táblát.) A mezőgazdaságban és erdőgazdálkodásban dolgozók korösszetétele tér el leginkább az átlagostól, ezen ágban igen alacsony a fiatalok és magas az idősek aránya. Természetesen a foglalkoztatottak többségét magukba foglaló szolgáltatás jellegű ágazatok részletesebb csoportjaiban nagyok lehetnek a különbségek (például a pénzügyi szektorban a fiatalok aránya nagyobb, mint a közigazgatásban dolgozók körében), ennek bemutatásával azonban itt nem foglalkozunk.

A foglalkoztatottak iskolázottság szerinti összetétele függ az adott nemzetgazdasági ág alapvető jellemzőitől, az ág jellegét döntően meghatározó tevékenység jellegétől, tartalmától. Az egyik végletet a mezőgazdaság és erdőgazdálkodás jelenti, ahol az általános iskola 8. osztályát végzetek vagy annál alacsonyabb végzettségűek még 2001-ben is a foglalkoztatottaknak több mint egyharmadát tették ki. A másik végletként a szolgáltatás jellegű ágak emelhetők ki, amelyekben a szellemi tevékenység dominál, nagyon magas a felsőfokú iskola legalább egy osztályát elvégzők aránya (26%).

#### TERÜLETI ADATOK

2001. évi népszámlálás képviseleti mintájának nagysága lehetővé teszi, hogy az országos adatok mellett nem csupán a hét régióra, hanem a megyékre, sőt a megyéken belül



még kisebb területi egységekre (az országgyűlési egyéni választókerületekre) vonatkozóan is rendelkezünk viszonylag megbízható információkkal. A továbbiakban a gazdasági aktivitással, a foglalkoztatottsággal, a foglalkoztatottak főbb strukturális jellemzőivel kapcsolatos mutatókat alapvetően a régiók szintjén vizsgáljuk, de utalunk a megyei sajátosságokra, sőt az egyéni választókerületekre is, amennyiben azok a nagyobb területi egységektől szignifikánsan eltérő és külön említésre érdemes sajátosságokat mutatnak.

A *foglalkoztatottság szintje* a 2001. évi népszámlálás időpontjában az országon belül meglehetősen jelentős területi eltéréseket mutatott. A legkedvezőbb (41,5%) Nyugat-Dunántúlon volt, több mint 5 százalékponttal haladva meg az országos átlagot. E térségen belül is kiemelkedőnek tekinthető a foglalkoztatottak aránya Vas megyében (43,2%). Az ország középső részén a népesség foglalkoztatottsági helyzetét elsősorban a fővárosi sokrétű munkalehetőségek határozták meg. Budapesten a népesség 42,2 százaléka volt foglalkoztatott, ami – az országos átlaghoz viszonyítva – 6 százalékpontos többletet jelentett. A közép-dunántúli térség, illetve az e térséghez tartozó megyék népességének foglalkoztatási mutatói szintén viszonylag előnyös elhelyezkedési lehetőségekről tanúskodnak.

A 2001. évi népszámlálás alkalmával egyértelműen két régió – Észak-Magyarország és Észak-Alföld – helyzete bizonyult a leghátrányosabbnak. Az előbbi térségben a népesség 30,9 százaléka, az utóbbiban 30,0 százaléka volt foglalkoztatott. Mindkét érték az országos átlag alatt maradt. Az e régiókhoz tartozó megyék közül Észak-Magyarországon Borsod-Abaúj-Zemplén megye, Észak-Alföldön Szabolcs-Szatmár-Bereg megye foglalkoztatottsági aránya (28,7, illetve 27,0 százalék) még az adott régió átlagánál is kedvezőtlenebb volt.

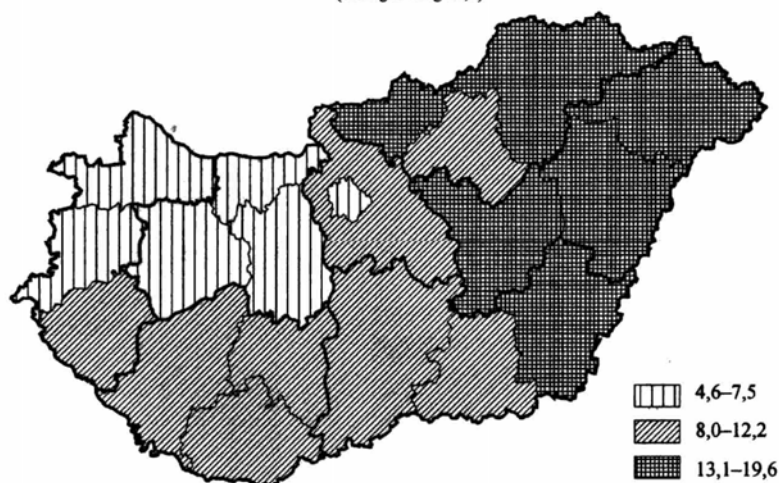
A munkanélküliség – érthető módon – viszonylag alacsony azokban a régiókban, ahol a foglalkoztatottság színvonala az átlagot meghaladja. Ez vonatkozik elsősorban Nyugat-Dunántúlra, továbbá Közép-Magyarországra és Közép-Dunántúlra. E területeken a munkanélküliek hányada 2,6–3,1 százalék közötti. Az északkeleti térségekben viszont az alacsony foglalkoztatottság – a munkanélküliek magasabb (6 százalék körüli) aránya mellett – azzal a következménnyel jár, hogy a népesség jelentős hányada kizorul a munkaerőpiacról, és az inaktív keresők vagy az eltartottak csoportjába tartozik.

A *munkanélküliségi ráták* vizsgálata a területi különbségeket még élesebben megvilágítja: e mutató értéke Nyugat-Dunántúlon alig háromötödét (5,8%), Közép-Magyarországon és Közép-Dunántúlon nem sokkal több mint kétharmadát (6,9, illetve 7,2 százalék) tette ki az országos átlagnak (10,2%). Észak-Magyarország és Észak-Alföld munkanélküliségi rátája viszont az országos szint másfélszeresét is meghaladta (16,4, illetve 16,3 százalék).

A nyugat-dunántúli régió legelőnyösebb helyzetű megyéjében – Vas megyében – a munkanélküliségi ráta 5 százalék alatt maradt (4,6%), az északi és az északkeleti térség leghátrányosabb helyzetű megyéiben viszont az országos átlag kétszeresét megközelíti: Borsod-Abaúj-Zemplén megyében 19,6, Szabolcs-Szatmár-Bereg megyében 19,1 százalék.

Még árnyaltabb képet kapunk, ha a vizsgálatba az országgyűlési egyéni választókerületeket is bevonjuk. Vas megye 4. számú választókerületében (Sárvár és környéke) a megyei aránynál is alacsonyabb a munkanélküliségi ráta (3,5%). A munkanélküliség szempontjából ugyancsak jó helyzetben levő Győr-Moson-Sopron megye 6. számú választókerületében (Kapunvár és környéke) a munkanélküliségi ráta értéke szintén 3,5 százalék, a megyei arány kétharmada.

1. ábra. A 15 éves és idősebb népesség munkanélküliségi rátája megyék szerint  
(országos átlag 10,2)



Fokozott figyelmet érdemelnek a magas arányú munkanélküliségű megyék különösen hátrányos helyzetű községei. Borsod-Abaúj-Zemplén megye 9. számú választókerülete (Encs és környéke) 31,8 százalékos munkanélküliségi rátája a megyei ráta másfélszeresét, az országos átlag háromszorosát is meghaladja. (E választókerületbe tartoznak többek között a megye nehezen megközelíthető kisebb települései és egyes határ menti települések.) A megyei szintet ugyancsak lényegesen meghaladják a 6. számú (Sajószentpéter és környéke), a 8. számú (Edelény és környéke), valamint a 12. számú választókerület (Tiszaújváros és környéke) 23-24 százalékos munkanélküliségi rátái. Bizonyos, hogy mindhárom körzetben jelentősen növelte a munkanélküliséget a nehézipar, és a bányák egy részének leépülése. Szabolcs-Szatmár-Bereg megyében a munkanélküliség különösen nagy mértékére utal a 6. számú (Nyírbátor és környéke), valamint a 8. számú választókerület (Vásárosnamény és környéke) rátája (28,7, illetve 27,7%), ezzel e választókerületek a hátrányos helyzetű megye fokozottan hátrányos helyzetű községeit képezik. Magas (20 százalékot meghaladó) munkanélküliségi ráták más – a munkanélküliség szempontjából kevésbé hátrányos helyzetben levő – megyék egyes községeiben is kimutathatók. Példaként említhető Békés megye 4. számú (Szeghalom és környéke), Hajdú-Bihar megye 4. számú (Hajdúhadház és környéke), 5. számú (Berettyóújfalu és környéke), 6. számú választókerülete (Püspökladány és környéke).

Néhány megyében, így például Bács-Kiskun megyében igen nagy szóródást lehet megfigyelni. A megye munkanélküliségi rátája 10,8 százalék, ami alig tér el az országos átlagtól. A megyeszékhely, Kecskemét két körzetében (1. és 2. számú választókerület) a ráta értéke csak 6,3, illetve 7,7 százalék, viszont a 10. számú választókerületben (Bácsalmás és környéke) 18,1 százalékos a munkanélküliségi ráta. E körzethez a megye déli részén fekvő, részben határszéli települések tartoznak. A megyén belüli legmagasabb ráta tehát közel háromszor akkora munkanélküliségre utal, mint a legalacsonyabb, és a megyei átlagnak kétszeresét is megközelíti.

Az elmondottak arra hívják fel a figyelmet, hogy a munkanélküliség vizsgálatánál nem elégedhetünk meg a régiókra vagy akár a megyékre vonatkozó információkkal, hanem feltétlenül szükség van mélyebb – kistérségi tartalmú – adatok ismeretére is.

A *foglalkoztatottak területi foglalkozási megoszlására* Budapest központi szerepe napjainkban is erősen rányomja bélyegét. A vezetők és az értelmiségi foglalkozásúak együttes aránya a 2001. évi népszámlálás időpontjában az országos 20,7 százalékkal szemben Közép-Magyarországon 27,1 százalékot, ezen belül Budapesten 31,9 százalékot, vagyis az országos átlag több mint másfélszeresét tette ki. A vezető állást betöltő, illetve a magasan kvalifikált dolgozók területi koncentrációjának mértékére jellemző, hogy míg az ország foglalkoztatott népességének egyötöde, addig a vezető, értelmiségi rétegbe tartozóknak közel egyharmada a fővárosban helyezkedett el. Az egyéb szellemi foglalkozásokba sorolt dolgozók foglalkoztatása Közép-Magyarországon, illetve Budapesten szintén meghaladta az átlagot. E dolgozók hányada országosan egyötöd, Közép-Magyarországon egynegyed, a fővárosban 26,7 százalék volt. A központi régiótól és a fővárostól eltekintve a vezetők, értelmiségiek, illetve az egyéb szellemi dolgozók hányada nem mutatott jelentős területi eltéréseket. Mindkét csoport részesedése az egyes régiókban egyötöd-egyhatod között volt.

A szolgáltatási tevékenységet folytatók aránya szinte valamennyi régióban 15-16 százalékot tett ki, egyedül Közép-Dunántúlon mutatott valamivel alacsonyabb értéket (13,7%).

A mezőgazdasági dolgozók aránya tekintetében az egyik végletet a Központi régió jelenti egyszázalékos részesedéssel, a másikat Dél-Alföld, ahol a mezőgazdasági munkát főtevékenységként folytatók hányada még mindig 8,6 százalék, vagyis az országos átlagnak közel háromszorosa. Az átlagosnál nagyobb még a mezőgazdasági tevékenység szerepe az Alföld északi részén, továbbá Dél-Dunántúlon. A Dunántúl középső és nyugati térségében, valamint Észak-Magyarországon a mezőgazdasági dolgozók aránya nem éri el a 3 százalékot.

Az ipari és építőipari tevékenységek jelentősége napjainkban főleg a Közép-dunántúli és Nyugat-dunántúli régióban érdemel figyelmet. E területeken a dolgozók közel kétötöde végez ipari, építőipari munkát. Az észak-magyarországi hagyományos iparvidéken ezek aránya (35,3%) az előbbieknél alacsonyabb, amit a nehézipar hanyatlása tesz érthetővé.

Az „egyéb” tevékenységet folytatók – elsősorban a segédmunkások – aránya nem mutat jellegzetes regionális eltéréseket: általában 8-10 százalék között ingadozik.

A területi egységekre is kiterjedő részletesebb elemzés lényegében arra a következtetésre vezet, hogy az adott terület jellegétől két fő réteg aránya függ erőteljesebben. Az egyik a vezető és értelmiségi rétegé, a másik pedig a mezőgazdasági dolgozóké. Az előbbieknél hányada elsősorban azokban a körzetekben kiemelkedő, amelyek egyetemi városokhoz, fontosabb kulturális, egészségügyi intézményekkel rendelkező településekhez, igazgatási központokhoz kapcsolódnak. Az utóbbiak továbbra is lényeges szerepet játszanak a hagyományosan mezőgazdasági jellegű körzetekben, ahol számottevő ipari beruházásokra, illetve a szolgáltatási intézmények hálózatának nagyobb arányú fejlesztésére nem került sor.

A vezető és értelmiségi foglalkozásúak aránya a 40 százalékot meghaladja, és ezzel az országos átlagnak kétszeresét teszi ki Borsod-Abaúj-Zemplén megye 2. számú választókerületében, amelybe kizárólag Miskolc város területe tartozik. E réteg a foglalkoztatottaknak szintén jelentős hányadát, közel egyharmadát (32,9%) alkotja Hajdú-

Bihar megye 3. számú választókerületében, amely Debrecen egyes részeire terjed ki. A vezető és értelmiségi foglalkozásúak részesedése több mint egyharmadot (35,0%) jelent Bács-Kiskun megye 2. számú választókerületében, amely a kecskeméti körzetek egyike. E réteg aránya szignifikánsan magasnak mondható még Somogy megye 1. számú választókerületében is (30%). A megyeszékhely, Kaposvár szavazókörének többségét ugyanis ebbe a választókerületbe sorolták.

Bizonyos, hogy több más megyeszékhelyen vagy kulturális centrumként ismert városban lehetne hasonló arányokat kimutatni. Utalni kell azonban arra, hogy a választókerületi beosztás nem vehette figyelembe a statisztika sajátos szempontjait. Ez természetesen a foglalkoztatottak társadalmi–foglalkozási rétegződésének a vizsgálatára is vonatkozik. A nagyjából egyöntetűnek mondható választókerületek mellett ezért „vegyes” jellegű választókerületek is kimutathatók, amelyekbe foglalkozási struktúra szempontjából eltérő városi és falusi területrészek együttesen tartoznak.

A mezőgazdasági foglalkozásúak arányát természetesen a földrajzi adottságok is számottevően befolyásolják. Érthető tehát a Dél-alföldi régió kiemelkedő szerepe a mezőgazdasági tevékenység szempontjából. Bács-Kiskun megyében az e csoportba tartozók aránya megközelíti a 10 százalékot, ami az országos átlag több mint háromszorosa. Egyes körzetekben a mezőgazdasági tevékenység még jelentősebb szerepet játszik: a 6. számú választókerületben például (Kiskörös és környéke) 22,7, az 5. számú választókerületben (Kiskunfélegyháza és környéke) 16 százalék a mezőgazdasági dolgozók hányada. Az ugyancsak tipikusan mezőgazdasági jellegű Csongrád megyében a 4. számú választókerületben (Csongrád és környéke) a foglalkoztatottaknak közel egyötöde (19,9%) mezőgazdasági foglalkozású, ami a megyei arálynak (8,7%) több mint kétszerese. Más területek földrajzi adottságai kevésbé kedveznek a mezőgazdasági tevékenységnek. Jellemző példa erre Borsod-Abaúj-Zemplén megye, ahol nem csupán a megyeszékhelyen és a többi ipari jellegű városban, hanem a megye más térségeiben is igen kicsi a mezőgazdasági tevékenységet folytatók foglalkoztatásban betöltött szerepe. Arányuk a választókerületek többségében 1-2 százalék körüli, vagy az egy százalékot sem éri el. (A megye átlaga 1,5 százalék.)

A nemzetgazdasági ág szerinti megoszlást területi szempontból vizsgálva megállapítható, hogy a mezőgazdaságban és erdőgazdálkodásban dolgozók regionális arányai lényegében azokat a jellegzetességeket mutatják, mint amelyeket a mezőgazdasági foglalkozásúakkal kapcsolatban megállapítottunk. A mezőgazdaságban és erdőgazdálkodásban főtevékenységet folytatók hányada legmagasabb a Dél-Alföldön (12,6%) és az Alföld északi területein, valamint a dél-dunántúli térségben (8,0, illetve 8,3 százalék) is meghaladja az országos átlagot.

Az iparba és az építőiparba sorolt ágazatok adatai – megerősítve a foglalkozási információkat – azt mutatják, hogy az iparosodottság szintje jelenleg a Dunántúl nyugati és középső térségében a legmagasabb. Az iparban és az építőiparban foglalkoztatottak együttes aránya a 2001. évi népszámlálás időpontjában Nyugat-Dunántúlon 41,1, Közép-Dunántúlon 43,1 százalékot tett ki. Ezek az országos arányt 8, illetve 10 százalékponttal haladták meg.

A gazdaság modernizációját egyértelműen igazolja, hogy a szolgáltatás jellegű ágak az ország valamennyi régiójában – a kevésbé fejlett keleti térségeket is ide értve – 2001-ben a dolgozók többségét foglalkoztatták. A szolgáltatás jellegű ágakban foglalkoztatot-

tak aránya Közép-Magyarországon (73%) megközelítette, ezen belül Budapesten (77,9%) meghaladta a dolgozók háromnegyedét.

A részletesebb területi vizsgálat a mezőgazdasággal és erdőgazdálkodással kapcsolatban megerősíti a régiók vonatkozásában elmondottakat. Azokban a körzetekben, ahol a mezőgazdasági foglalkozásúak aránya kimagasló, a megfelelő nemzetgazdasági ág aránya is kiemelkedik más körzetek közül.

Az iparban és építőiparban foglalkoztatottak részesedése általában a Dunántúl erősen iparosodott térségeihez tartozó körzetekben sem éri el az 50 százalékot. Kivételt képez Komárom-Esztergom megye 3. számú (Kisbér és környéke), továbbá 4. számú (Komárom és környéke), Veszprém megye 1. számú (Ajka és környéke), valamint 5. számú választókerülete (Várpalota és környéke). E Komárom-Esztergom és Veszprém megyei körzetekben a dolgozók többségét (mintegy 51-55 százalékát) az iparban és az építőiparban foglalkoztatják. A korábban ipari jellegűnek tekintett Észak-magyarországi régióon belül jelenleg csak egy körzetben, Borsod-Abaúj-Zemplén megye 7. számú választókerületében (Kazincbarcika és környéke) éri el az iparban és építőiparban foglalkoztatottak hányada az 50 százalékot.

A választókerületi részletezettségű adatok is bizonyítják, hogy az ország legtöbb körzetében érvényesül a szolgáltatás jellegű ágak elsőbbsége. Csak elszórtan található olyan körzetek, ahol a szolgáltatás jellegű ágazatokban foglalkoztatottak nincsenek többségben, de arányuk e körzetekben is meghaladja a 40 százalékot.

### A KÖZÉP-MAGYARORSZÁGI RÉGIÓ

A foglalkoztatottsági adatok megyénkénti és régiókénti különbségei jól jelzik, hogy hazánkban még kis területein is jelentős, földrajzilag is körülhatárolható eltérések tapasztalhatók, de ha az adatok megyén belüli differenciáltságát is megvizsgáljuk, további különbségek is felfedhetők. A magyar népszámlálások történetében egyedülálló módon, a teljes körű adatállományok nyilvánosságra hozása előtt már olyan adat-összeállítás is megjelent – a választókerületek adatait tartalmazó, már említett kiadványról van szó –, amely alkalmas a kisebb területi egységek foglalkoztatottsági adatainak az elemzésére is. Tekintettel a Közép-magyarországi régió kiemelt jelentőségére – a népesség 28 százaléka, a foglalkoztatottak 32 százaléka él ebben a régióban – a kisebb területi egységek szerinti vizsgálatot nem csak kiegészítő jelleggel alkalmazzuk, hanem a hangsúlyt ezekre a vizsgálatokra helyezzük.

A Közép-magyarországi régió a fővárosból és Pest megyéből áll. Pest megye az ország sajátos megyéje, központja egy másik „megyében”, Budapesten van, területét a Duna vágja ketté, egyes részei – a budapesti agglomerációhoz tartozó települések – rendkívül szorosan kötődnek a fővároshoz, más részei pedig a szomszédos megyeközpontok felé gravitálnak. Igazi egységet tulajdonképpen a főváros és vonzáskörzete képez, ha a megyerendszer nem lenne szigorúan szabályozott, akkor a Közép-magyarországi régió ebből a területegyüttesből kellene állnia. Az agglomerációhoz tartozó Pest megyén belüli településegyüttest nem lehet külön vizsgálni, mert a választókerületek nem e szempontok szerint lettek kialakítva, azt azonban tudjuk jelezni, ha egy választókerületben az agglomerációs települések többségben vannak (ilyen például a 8. számú választókerületben Érd, Diósd, Százhalombatta).

A régió természetes központja Budapest, mely az ország legfontosabb gazdasági-közigazgatási centruma. Magyarország is azokhoz az európai országokhoz tartozik, melyeknek életében egy város, többnyire a főváros az ország lélekszámához képest kiemelt jelentőségű. Budapest területe is viszonylag kiterjedt, földrajzi fekvése is szerteágazó, már csak azért is mert síkvidék és hegyvidék találkozási pontján fekszik, és területét Európa legnagyobb folyója szeli ketté. Ezt is figyelembe véve ahhoz, hogy a rendelkezésünkre álló információkat elemezni és megfelelőképpen szemléltetni tudjuk, különböző kerület- (Pest megyében település-) csoportokat képeztünk. (Bár egyes választókerületek átlépik a kerülethatárokat, de néhány kivételtől eltekintve a főbb kerületcsoportok adatai összeállíthatók.) Az elemzés során – a választókerületi adatok segítségével – az egyes kerületeken belüli jellemző különbségeket is bemutatjuk, melyek arra utalnak, hogy Budapesten is megtalálhatók a túlszűfolt, ma már sok szempontból elavult belső városrészek, a lakóházakkal vegyesen telepített, gyárnegyedekből álló, a nagyvárosi színvonalnál elmaradó, viszonylag rendezett övezetek, a viszonylag jelentős területet felölelő villanegyedek és a nagy kiterjedésű, családi házas, helyenként falusias, külterjes, hézagosan beépített övezetek.

A 2001. évi népszámlálás előzetes adatai szerint Budapest népessége 1990 óta jelentősen csökkent. A csökkenés mértéke Budapesten belül területenként különböző mértékű volt, de ennek vizsgálatától a jelen értékelésnél eltekintünk, tekintettel arra, hogy minden részletre kiterjedő elemzést csak a teljes körű adatállomány alapján érdemes elvégezni. A 2001. évi előzetes adatok is jól jelzik, hogy a főváros egyes területei lakosságának szerkezetéről eltérő. (Lásd az 5. táblát.)

5. tábla

*A népesség és a foglalkoztatottak főbb jellemzői kerületcsoportonként*

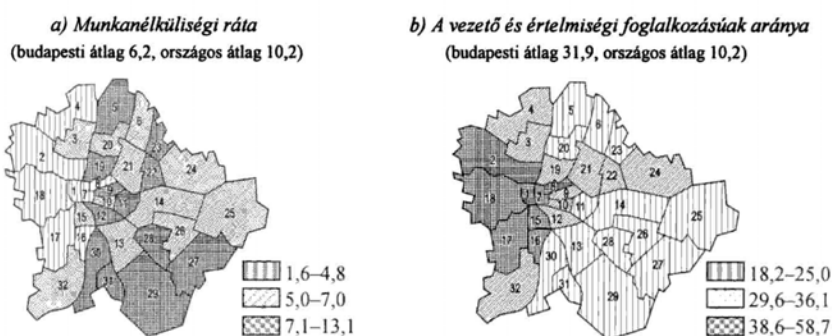
Terület	0–14	60–	Felsőfokú iskola 1 vagy magasabb évfolyama a 25 éves és idősebb népesség	A foglalkoztatottak aránya a 15 éves és idősebb népesség	A gazdaságilag aktív népesség aránya a 15 éves és idősebb népesség	A vezető, értelmi-ségi foglalkozásúak aránya a foglalkoztatottak	Szolgáltatási jellegű ágban dolgozók aránya	A munkanélküliségi ráta (százalék)	
	éves népesség aránya (százalék)		százalékában						
Ország	16,8	20,2	13,5	43,6	48,5	20,7	61,3	10,2	
Budapest	11,5	18,9	25,7	48,7	52,0	31,9	77,9	6,2	
Észak-Buda (I., II., III., XII. ker.)	11,6	25,8	42,0	49,2	51,0	46,2	81,3	3,6	
Dél-Buda (XI., XXII. ker.)	13,0	23,2	34,6	47,5	50,0	41,3	79,3	5,1	
Buda együtt	12,2	24,8	39,1	48,5	50,6	44,4	80,5	4,2	
Észak-Pest (IV., XIII., XV. ker.)	14,7	18,6	20,4	51,8	55,7	24,2	77,8	6,9	
Kelet-Pest (X., XIV., XVI–XIX. ker.)	13,4	21,1	19,7	49,0	52,6	25,8	75,2	6,8	
Dél-Pest (IX., XX., XXI., XXIII. ker.)	14,1	21,9	13,2	47,5	51,4	24,2	75,5	7,6	
Belső-Pest (V.–VIII. ker.)*	12,3	23,7	28,0	46,1	49,7	35,6	80,5	7,2	
Pest együtt	11,3	16,6	20,3	48,8	52,5	26,9	76,9	7,0	

\* A XIII. kerület egy kis területi egységével együtt.

A budai kerületek lakossága jóval idősebb a Duna bal parti kerületeiben élőkénél. Természetesen a kerületcsoportokon belül is nagy különbségek lehetnek, például Észak-Budán a XII. kerületben lakók 31, a III. kerületben lakók 21 százaléka tartozott a 60 éves és idősebbek közé. Ha ugyanebben a kerületcsoportban választókerületenként vizsgáljuk az adatokat, akkor még nagyobb különbségeket tapasztalhatunk, például az 1-es választókerületben (Vár és környéke a II. kerület egy részével) az időskorúak aránya 32 százaléka, a 4-es választókerületben (Római fürdő, Csillaghegy, Békásmegyér) csak 15 százaléka.

Budapesten a munkanélküliségi ráta az országos átlaghoz viszonyítva alacsonynak mondható, de ez a városon belül viszonylag nagy különbségeket takar. Ebben a tekintetben is nagy eltérés van a budai és a pesti oldal között: az előbbi területen 4,2, az utóbbi térségben 7 százalékos a munkanélküliség. Mind a budai, mind a pesti kerületcsoportokon belül is jelentősek a különbségek. A választókerületek szerinti adatok igen nagy különbséget mutatnak: Budán 1,6-től (2. választókerület Hűvösvölgy, Pesthidegkút) 5,3 százalékgig (3. választókerület Flórián tér környéke), Pesten pedig 4,1-től (7. választókerület Belváros, Lipótváros) 13,1 százalékgig (11. választókerület Józsefváros, Tisztviselőtelep) terjed a skála. (Lásd a 2. ábra a) kartogramját.) Az egyes társadalmi rétegek munkanélküliségének mértéke jelentősen különbözik, ezért a Budapesten belüli eltérések összefüggésben vannak azzal, hogy a magasabb vagy alacsonyabb státusú lakossági rétegek a főváros mely területén laknak. (A munkanélküliségi ráta ugyanis a magasabb státusú népesség körében az átlagosnál jóval alacsonyabb.)

2. ábra. A budapesti 15 éves és idősebb népesség választókerületek szerinti



A népszámlálási felvételek hagyományosan nem tartalmazzák a jövedelemre vonatkozó kérdéseket, ezért a társadalmi rétegződési vizsgálatokra csak közvetett információk alapján van lehetőség. Ezek közé tartozik az iskolai végzettség és a foglalkozás kérdésköre. Az egyszerűség kedvéért most csak a foglalkozással kapcsolatos információkat vizsgáljuk, ugyanis a foglalkoztatottak foglalkozási összetétele összefüggést mutat az adott területen élő lakosság iskolai végzettség szerinti megoszlásával. A budai, illetve a pesti kerületekben élő foglalkoztatottak foglalkozás szerinti összetétele lényegesen eltér egymástól: Budán a vezető értelmiségi foglalkozásúak aránya 44,4, Pesten 26,9 százalék volt. (A 25 éves és idősebb népesség körében a felsőfokú iskola 1 vagy magasabb évfo-

lyamát végzők aránya a két városrészben 39,1, illetve 20,3 százalék.) Így a más típusú foglalkozást folytatók aránya is kisebb-nagyobb mértékben eltér, például a segédmunkások aránya Budán 4,7, Pesten 8,3 százalék volt. (A magasabb iskolai végzettséget kevésbé igénylő egyéb szellemi foglalkozásúak foglalkoztatottakon belüli arányában viszont kisebb – 25,4, illetve 27,2 százalék – az eltérés a két városrész között.)

Ha az összes szellemi foglalkozású személyt egy kategóriának vesszük, akkor egyes budai kerületekben (választókerületekben) a koncentráció mértéke igen erőteljes. Például a XII. kerületben (18. választókerület) a szellemi foglalkozásúak aránya meghaladja a 83 százalékot, az I. és a II. kerület egy részében (1. választókerület a Vár és környéke) pedig majdnem eléri a 80 százalékot. Az országos átlag 40,7, a budapesti átlag pedig 58,6 százalék.) Ha választókerületi szinten elemezzük az adatokat, akkor Budán csak egy olyan választókerület van, ahol a szellemi foglalkozásúak aránya 60 százaléknál alacsonyabb. A budai és a pesti városrészek foglalkoztatottjai között a szellemi, illetve a fizikai foglalkozásúak megoszlását illetően mindig is volt különbség, ez az eltérés azonban népszámlálásról népszámlálásra fokozatosan nőtt. A pesti kerületekben a szellemi foglalkozásúak koncentrációja – alacsonyabb szinten – szintén számottevő. A belső városrészekben az V. kerület és a XIII. kerület egy részének (7. választókerület Belváros, Lipótváros) és a VI. kerületnek (8. választókerület) foglalkoztatottjai között a szellemi foglalkozásúak aránya 70 százalék körül alakult, Csepelen a 30. választókerületben ugyanez a mutató 44,2 százalék volt. (Pesten itt volt a legalacsonyabb a szellemi foglalkozásúak aránya, de még az is az országos átlag felett van.) A pesti városrész egyes választókerületeiben viszont – a budapesti, illetve az országos átlaghoz képest is – magas a segédmunkások aránya. Például a IV. kerület 5. választókerületében (Újpest, Káposztásmegyér) a foglalkoztatottaknak 13,2 százaléka (a férfiaknak 16 százaléka), a XIII. kerület 19. választókerületében (Váci út, Lehel út, Lehel tér) 12,8 százaléka (a férfiaknak 14,7 százaléka) volt segédmunkás. E néhány adat is jól mutatja, hogy a Budapesten élő népesség társadalmi–foglalkozási összetétele területenként mennyire eltérő, és az elkülönültség kisebb területi egységek felé haladva milyen jelentős értéket mutat. Ezt jelzi a 2. ábra *b*) kartogramja is, amely a budapesti vezető és értelmiségi foglalkozásúak választókerületenkénti arányát mutatja.

A lakosság demográfiai, foglalkozási jellemzői a Közép-magyarországi régió Pest megyei választókerületeiben is jelentősen eltérnek egymástól, bár a különbségek nem annyira szélsőségesek, mint Budapesten. Pest megye 16 választókerületéből négy választókerület (8., 9., 10., 11.) a budai, a többi a pesti oldalon (Csepel szigeten) található. A két oldal választókerületeiben élő lakosság korösszetétele nem tér el jelentősen egymástól, inkább egyes választókerületek között vannak nagyobb eltérések. Például a 9. választókerületben (Budaörs és környéke) a 60 éves és idősebbek aránya 16,3, az 5. választókerületben (Aszód és környéke) 21,9 százalék.

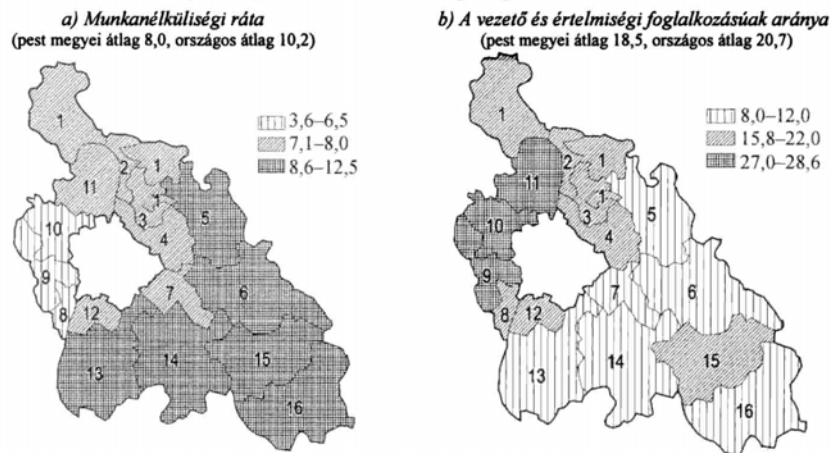
A munkanélküliség tekintetében a Duna két partja között az eltérések már jelentősebbek, a budai oldalon a munkanélküliségi ráta 6 százalék – ezen belül a 10. választókerületben (Pilisvörösvár és környéke) 3,6 százalék –, a pesti oldalon 8,8 – ezen belül a 14. választókerületben (Dabas és környéke) 12,5 – százalék volt. (Lásd a 3. ábra *a*) kartogramját.)

Az iskolai végzettség vonatkozásában is jelentősek az eltérések a budai és a pesti oldal között. A budai oldal négy választókerülete 25 éves és idősebb népességének 19 százaléka végzett legalább egy évfolyamot felsőfokú iskolában, a pesti oldal 12 választókerületében ezek aránya csak 9 százalék volt. Ez utóbbi térségben ez az arány a 2. választó-



kerületben (Vác és környéke) volt a legmagasabb (16%) és a 6. választókerületben (Nagykátá és környéke) a legalacsonyabb (4%). Hasonlók az eltérések a foglalkozási összetételben. A budai oldal választókerületeiben a foglalkoztatottak között a szellemi foglalkozásúak aránya 48,4, a pesti oldal választókerületeiben 35,3 százalék volt. A budai oldal foglalkoztatottjainak foglalkozási összetételére leginkább a pesti oldal északi részén levő három választókerületben lakók (1. választókerület Szob és környéke; 2. választókerület Vác és környéke; 3. választókerület Dunakeszi és környéke) foglalkozási összetétele hasonlít, ugyanis ezen a területen a szellemi foglalkozásúak aránya közel 44 százalék volt. Kelet-Pest négy választókerületében viszont ez a mutató csak 33 százalék. (Lásd a 3. ábra b) kartogramját.)

3. ábra. A pest megyei 15 éves és idősebb népesség választókerületek szerint



Összefoglalóan azt lehet mondani, hogy Pest megyében a budapestihez hasonló helyzetet figyelhetünk meg. A megye kedvezőbb adottságú hegyekkel, dombokkal tagolt és a Duna közeli területein a magasabb társadalmi statusú rétegek élnek, a keleti, déli síkvidéki területeken pedig a hátrányosabb helyzetű lakossági csoportok találhatók. Bár vándorlási adatok jelenleg még nem állnak rendelkezésünkre, de feltételezhető, hogy az elmúlt évtizedben Budapestről kiköltözött több mint 100 ezer ember főleg Pest megye budai oldalán, illetve a pesti oldal északi településein telepedett le.

#### SUMMARY

The study shows the main characteristics of the labour market situation on the basis of the information coming from data of representative sample of the 2001 census. In order to correct interpretation it provides a review of the antecedents of the census, pilot censuses, of the questions relating to employment of the final questionnaire and experiences of enumeration and processing. The study provides analysis on employment, unemployment and the main structural characteristics of the employed. Sample size makes it possible to analyse and show the characteristics of the different regions and counties, besides national data. Since the most important results of the census were processed for individual election districts (in order to meet requirements of the preparatory work relating to the 2002 elections), the authors show some special characteristics of these smaller spatial units.

## SZEMLE

---

### ÉVNYITÓ ÉRTEKEZLET A KÖZPONTI STATISZTIKAI HIVATALBAN

2002. január 22-én a korábbi évek hagyományainak megfelelően, a Központi Statisztikai Hivatal elnöksége megtartotta összevont vezetői értekezletét, amelynek előadója ezúttal is *Mellár Tamás*, a KSH elnöke volt.

Bevezetőjében *Balthasar Gracián y Morales* spanyol írótól kölcsönzött gondolattal arra utalt, hogy gyakran a rendkívüli helyzetek, a cselekvési kényszerek azok, amelyek ki tudják váltani a cselekvő ember legjobb tulajdonságait. Mindezt azért mondta, mert bár a 2001-es év minden szempontból igen sikeres volt a KSH számára, 2002-ben azonban olyan nehéz körülmények elé nézünk, hogy nem lehet a régi módon cselekedni, mert akkor megindul a lassú leépülés. Ahhoz, hogy a KSH képes legyen megőrizni elért eredményeit, pozícióját, tekintélyét, reformértékű átalakulást kell kezdeni már az év elején. Ez ad különös jelentőséget az idei értekezletnek.

Az előadó négy téma köré építette fel előadását. Ezek a következők voltak:

- az elmúlt év gazdálkodása anyagi–műszaki eredményei, helyzete;
- a 2001-es év szakmai eredményei és hiányosságai;
- főosztályi körkép;
- a jövő feladatai.

A témák részletes kifejtése a következő volt.

1. *Az elmúlt év gazdálkodása*, anyagi–műszaki feltételei igen kedvezőek voltak. Az igazgatóságokon átlagosan 60, a központban 50 százalék volt a bérek növekedése, és így az év végére az átlagbér a megyéknél 144 ezer, a központban 184 ezer forint volt. A Hivatal dolgozói átlagosan 3 és félhavi jutalmat kaptak, és 2001-ben két alkalommal kaptak 13. havi fizetést. A személyi javadalmazásokon túl teltett eszközök felújítására is, aminek egyik legfontosabb elemeként az összesen 150 millió forint értékű bútorvásárlást említette az előadó. Ami a beruházásokat illeti, a Szolnok Megyei Igazgatóság épületfelújítása befejeződött, a baranyai és a csong-

rádi felújítás elkezdődött, emellett folyik a központi épület homlokzatának felújítása és megújult a KSH nagy tanácsterme is.

2002-re azonban a KSH költségvetése a tavalyinál jóval szerényebb: 20,2 helyett 10,5 milliárd forint lesz. Ezért a várható bruttó bérnövekedést 7,8 százalékra, az év közben kiosztható jutalmat 1 hónapra tervezik. Figyelembe véve a kötelező létszámleépítést a KSH létszáma a tárgyévben 34 fővel bővíülhet. Az elnök ennek kapcsán megjegyezte, hogy ez valószínűleg nem jelent azonnal főosztályokra leosztható létszámot, hiszen célszerű lenne ennek nagy részét fiatalok beállítására fordítani, akiket belépésükkor egy központi keretbe sorolnának, majd megfelelő kiképzés, és néhány hónapi rotálás után kerülhetnének végleges főosztályi helyükre. A beruházásokra és felújításokra erre az évre mintegy 1 milliárd forintot tervez a vezetés, aminek nagy része a számítógép-állomány cseréjére és felújítására szolgál, de ennek terhére kell befejezni, illetőleg folytatni a megkezdett épület-felújításokat is.

2. Az elmúlt év *szakmai értékelésére* áttérve az előadó először a kedvező jelenségeket, majd a hiányosságokat vette sorra.

A *pozitívumokat* részletezve, a következőkre helyezte a hangsúlyt.

– A Hivatal múlt évi tevékenységéből kiemelkedett a népszámlálás, amely nehezen indult, sok előzetes félelemmel kellett megküzdeni a végrehajtással kapcsolatban, de ezek szerencsére alaptalanul bizonyultak. A siker három fő momentumra közül az első az adatfelvétel, melynek során a megyék az önkormányzatokkal együttműködve igen jó munkát végeztek. Kiemelte azt, hogy különösen nagy feladat hárult a fővárosi igazgatóságra, amelyik e téren élenjáró munkát végzett.

– A második említésre érdemes mozzanat az OCR-es feldolgozás sikere volt. Ez a fázis határidő előtt, simán, különösebb gond nélkül lezajlott, így méltán lehetünk büszkéek arra, hogy a Bull cég ezt a

munkát azóta referenciamunkának tekinti. Végül a népszámlálás harmadik sikeres eleme a minta gyors és jó minőségű feldolgozása volt, hiszen december 12-ére készen állt az első olyan adatközlés, amely lehetőséget adott egy nagy érdeklődést kiváltó sajtó-értekezletre. Ez a gyors munka a Népszámlálási és az Informatikai főosztályok hatékony együttműködésének jó példája volt.

– A másik sikeres akció a Szőlő- és gyümölcs-összeírás volt. Ezzel kapcsolatban mintegy példaként azt említette az előadó, hogy a szőlőültetvény összes területére vonatkozó, és az adminisztratív nyilvántartásoknak lényegesen ellentmondó eredményt a szakma azonnal, és kételkedés nélkül elfogadta. Ez a felvétel azért is jelentős, mert ezzel Magyarország a tagjelölt országok közül elsőnek teljesítette a három nagy összeírásra vonatkozó EU-elvárást.

– A harmadik fontos esemény a Hivatal életében a kanadai szakértői átvilágítás volt. Ennek legfontosabb eredménye az volt, hogy a Hivatal jó értékelést kapott, amely értékelésből ezúttal talán legfontosabb a KSH politikamentes magatartásának hangsúlyozása. Emellett a jelentés egy sor javaslatot is ad a munka jobbítására, amely javaslatokat nekünk kell értékelnünk, értelmeznünk, és ha úgy találjuk, megvalósítanunk. A javaslatok értékelésére létrehozott öt munkacsoport már megkezdte tevékenységét.

– Lényeges haladás történt az adattárház és a regisztermunkák területén, ami azzal biztat, hogy márciusra jól működő adattárház üzemel, a statisztikai regiszter pedig hozzájárul az adatminőség javításához.

– A kiemelt feladatok árnyékában csendben, rendben folytak az EU-csatlakozásra való felkészülés munkái, a harmonizációk. Ez a munka nagy nemzetközi elismerést vívott ki, amit számos találkozó, tanfolyam és egyéb rendezvény is bizonyít.

– Ugyancsak a pozitívumok között említette az előadó azt, hogy folyamatban van a könyvtár új számítógépes rendszerének elkészítése.

– Az említett sikeres munkák és kedvező jelenségek mellett volt még egy sor olyan esemény, amely összességében pozitívan értékelhető, ám valamennyi esetben valami fenntartást is meg kell említeni. Ezek közül csak mintegy példaként említette az elnök a Magyar Nemzeti Bankkal létrejött megállapodást, ami jó, de féltő, hogy alacsonyabb szinteken, a részletek még sok súrlódásra adnak okot. Példaként említette a holland projekt beindulását is, ami személyi és technikai problémák miatt a vártnál lassabban halad. Végül itt emelte ki a kitűnő megyei évkönyveket, ám ezek némelyike csak később, 2001. november derekán jelent meg.

A kedvező eredmények és jelenségek után az előadó rátért a *negatívumokra*, melyek – megállapítása szerint – jóval kisebb számban fordultak elő.

– Ezek közül első helyen emelte ki azt, hogy a központ és a megyék kapcsolata továbbra sem zökkenőmentes. Nehézkes a gépezet, akadozik a kommunikáció, és ezt csak súlyosbítja az a tény, hogy a központon belül is vannak információáramlási zavarok. A régi döntéselőkészítő mechanizmus megszűnt, az új még nem működik, és hiányzik a vezetői kollégium intézménye. Az elnökség elszánt abban a kérdésben, hogy a döntési mechanizmust a szakmai egyeztetések rendszerén keresztül megszilárdítsa. Az előadó három kis példát említett a működési zavarokra. Az első a kanadai átvilágítással volt kapcsolatos: az elkészült jelentés minden ígéret és meg egyezés ellenére hosszú ideig nem volt megtalálható sem az intraneten, sem pedig az interneten. A másik példa a népszámlálás december 12-ei első eredményeinek bemutatásáról szól: az egyébként igen sikeres munka értékét nagyban rontotta, hogy ezeket az eredményeket nem megfelelőképpen terjesztették, és még január közepén is volt olyan megye, ahová nem jutottak el a megfelelő anyagok. A harmadik történet szerint az egyik megyei igazgatóság négy napig nem működött a számítógépes hálózat, ám erről a KSH elnökségének egyik tagja sem értesült.

– Negatívumként említette az elnök az internetes megjelenés, elsősorban az összehangoltság és a koncepció hiányosságait.

– Rossznak tartotta a KSH és a külvilág kapcsolatát. Jóllehet a sajtókapcsolat jól működik, az év sajtócsapata valóban elismerésre méltó munkát végzett, még sincs e téren igazi előrelépés. Az előadó úgy fogalmazott, hogy elébe kell menni a kritikáknak, ezért a kanadai jelentést értékelő egyik bizottság feladata éppen a Hivatal és a külvilág kapcsolatainak rendezése lesz.

3. Az értekezlet harmadik napirendi pontjaként a *főosztályi körképre* került sor, aminek során az elnök tömören értékelt az egyes főosztályok és önálló osztályok elmúlt évi teljesítményét.

– Az elnöki felügyelet alá tartozó blokkból először a *Nemzetközi főosztályt* említette. Ez a főosztály az elmúlt évben igen sok feladatot kapott a kanadai átvilágítás, a PHARE, valamint a kétoldalú megbeszélések és a csatlakozás előkészítése kapcsán. A főosztály ezeket a feladatokat jó minőségben és határidőre elvégezte. A *Statisztikai Szemle* továbbra is megbízhatóan, kiegyensúlyozottan működik. A megjelent anyagok színvonala azonban nem nőtt, sőt a publikációs kedv lanyhulásával inkább csökkenő színvonalat lehet prognosztizálni. Ezért – hangsúlyozottan nem anyagi okokból – felvetette azt a kér

dést, hogy a Hivatalnak szüksége van-e négy folyóiratra. Az *Ellenőrzési osztály* munkájával korábban nem volt elégedett, ám a múlt évben némi jó irányú elmozdulás volt tapasztalható. A *Módszertani osztály* még túlságosan új ahhoz, hogy érdemi eredményeket lehessen várni tőle. Az a tény, hogy létrejött, mindenesetre öröndetes, mint ahogy biztatók első feladatában, a szezonális kiigazító eljárások egységes rendszerének kialakításában eddig elért eredményei is.

– A koordinációs elnökhelyettes felügyelete alá tartozó főosztályok közül a *Területi és Koordinációs főosztály* nagy rutinnal végzi munkáját. Munkájukban nincs áttörés, ám az internetes adatgyűjtés és a regisztermunkák terén jó eredményeket értek el. A *Tájékoztatási főosztály* élén vezetőváltás volt, ami várhatóan még jobb teljesítményhez vezet. A tájékoztatói koncepció, a sajtószoza, a nagyszámú tájékoztató impozáns, ám a külvilággal való kapcsolatban vannak hiányosságok. Az *ECOSTAT* tudott élni helyzeti előnyével, stabilizálódott, és igen találatos. Úgy tűnik, hogy sikerült valóban megtalálni saját profilját. A *Kommunikációs és Marketing főosztály* létrejöttét az indokolta, hogy nem volt átfogó gazdája a teljes kommunikációnak. A jövőben igen fontos feladatok elé néznek.

– A társadalomstatistikai ágon a *Népesedés-, Egészségügyi és Szociális Statisztikai főosztályon* a sikeres kiadványok (Demográfiai évkönyv, Vándorlásstatistika stb.) jelzik a helyes utat és a vezetőváltás eredményességét. A *Társadalomstatistikai főosztály* az időmérleg-felvétel eredményeit most elemzi és publikálja, de kérdés, mi lesz ezután. Az *Életszínvonal- és Emberierőforrás-statisztikai főosztály* folyamatosan jól dolgozik, bár gondok voltak az oktatásstatistikával. A jövőben nehéz feladatok várnak a főosztályra. A *Népszámlálási főosztály* az év legkellemesebb meglepetését okozta a népszámlálás sikeres lebonyolításával, amihez az elnök külön gratulált. A *Környezetstatistikai főosztály* csendben, rendben végzi munkáját, amit tennie kell, megteszi, időre elkészül feladataival. A *Népességtudományi Kutatóintézet* presztízse nőtt, az intézmény magára talált. Sikeres munkák sora igazolja a vezetőváltást.

– A gazdaságstatistikai blokkból a *Nemzeti Számlák főosztálya* őrzi elért eredményeit. Bár áttörésjellegű változás nincs, a negyedéves GDP-számítások jó irányban haladnak. Az *árstatistika* kiemelkedően stabil, pontos, jó minőségű. A *Mezőgazdasági főosztály* jól szervezett csapatmunkát végzett az elmúlt évben is, a szőlő- és gyümölcsfeldvétel sikeres végrehajtása nem volt váratlan eredmény. Az *Iparstatistikai főosztály* szintén magas színvonalú munkát végzett. Kiemelkedő szerepe volt

az EU harmonizációs feladatok teljesítésében és az adattárház kezdeményezésében. A *Szolgáltatásstatistikai főosztály* munkája lényegesen nem változott. Újabb kísérleteket tesz a kiskereskedelmi statisztika javítására, az információstatisztika fejlesztésére, de a korábbi problémák tovább élnek. A *Külkereskedelm-statisztikai főosztály* hosszú vajúdas után jött létre, nehéz indulása volt, de esélyük van a jó folytatásra. Fontos feladatok várnak erre a főosztályra az EU-konform külkereskedelmi statisztika kialakításában. A *Pénzügystatistikai főosztály* jól teljesített, bár feladata nem volt könnyű. A jövőben munkájuk talán legnagyobb kihívását a nehezen kezelhető partnerek fogják jelenteni.

– Az ellátó és adminisztrációs blokkon belül az új *Költségvetési és Pénzügyi főosztály* működésének legnagyobb dicsérete az, hogy munkájukat nem vesszük észre. Új feladatuk, a főosztályi szintű gazdálkodás megszervezése és lebonyolítása, nehéznek ígérkezik. A *Műszaki és Ellátási főosztály* tevékenységével kapcsolatban sok probléma adódott, ami a közeljövőben vezetőváltást, és új működési stílust indokol. Az *Informatikai főosztály* kiemelkedően jó teljesítményt nyújtott, kulcsszerepe volt a Hivatal sikereiben. Átszervezését a legkevésbé sem a hiányosságok indokolják, hanem a már-már kezelhetetlenül nagy mérete.

4. Utolsó témaként az előadó a 2002. év feladatait tekintette át.

– Ezen feladatok közül elsőnek említette a szervezeti és működési rend javítását. Be kell fejezni a döntési mechanizmus véglegesítését. Ennek keretében kiemelt feladat a belső tervezés és költségelszámolás átalakítása. Ugyancsak megkezdődik megyei igazgatóságok helyének és szerepének átgondolása; ezek működésével kapcsolatban is alapos átalakításra lesz szükség. Tekintve, hogy külső forrás nem áll rendelkezésre, mindezeket a feladatokat belső erőből kell megvalósítani.

– A kiemelt fontosságú feladatok között említette az elnök az elektronikus adatgyűjtés fejlesztését, a kistérségi rendszerek felülvizsgálatát, valamint azt, hogy foglalkozni kell az államigazgatási statisztika szerepével, a külvilággal fenntartott kapcsolat jobbításával, és javítani kell a személyzeti és oktatáspolitikát.

– A kisebb, ám el nem hanyagolható feladatok közül kiemelte a nemzeti számlák továbbfejlesztését (elsősorban a forrás/felhasználás táblák szélesebb körű alkalmazását, valamint a negyedéves GDP szezonális kiigazítását), az Intrastat-rendszer kiépítésének folytatását, valamint azt, hogy az EU csatlakozási tárgyalások lezárásáig még sok jelentést kell készíteni, sok rutinmunkát kell elvégezni. Meg kell

reformálni a lakossági adatgyűjtést, a gyorstájékoztatókat, és végül létre kell hozni egy megújult, egységes szemléletű Szervezeti és Működési Szabályzatot.

Az elnök azzal a *Széchenyitől* kölcsönzött gondolattal fejezte be előadását, miszerint a magyar nemzet egyaránt hajlamos mind az alaptalan borúlátásra, mind a hasonlóképpen megalapozatlan derűre, holott ahhoz, hogy nagy eredményeket érjünk el, csak a dolgok reális szemléletén és értékelésén keresztül vezet az út. Ezért, az elért eredményekre támaszkodva, ám a jövő nehézségeinek tudatában kell munkálkodnunk.

Az előadást követően lehetőség volt kérdések és ellenvélemények megfogalmazására, vitára, ám

mindössze egy hozzászólás volt: a Külkereskedelem-statisztikai főosztály vezetője kifogásolta az új főosztály elhelyezési körülményeit, és kérte azok javítását. Ugyancsak ő tette szóvá azt, hogy a KSH adatainak csak kis hányada férhető hozzá ingyenesen, míg más európai hivatalok ennél jóval szélesebb körben biztosítják ezt. Mellár Tamás válaszában egyrészt ígéretet tett a munkakörülmények javítására, másrészt emlékeztetett arra, hogy más hivatalok (például a kanadai hivatal) finanszírozása az itthoniól eltérő struktúrájú, és ez ad lehetőséget arra, hogy viszonylag nagyobb adatkört tudjanak ingyenesen a felhasználók rendelkezésére bocsátani.

H. L.

## A STATISZTIKA HETE BÉCSBEN

Az Osztrák Statisztikai Hivatal (Statistik Austria) és az Osztrák Statisztikai Társaság nemzetközi konferenciája Bécsben került megrendezésre. A nemzetközi konferencia, amelyet 2001 október 16. és 19. között tartottak, az 50 éves Osztrák Statisztikai Társaságot ünnepelte. A két fő szervező mellett támogatója volt a rendezvénynek az Osztrák Szövetségi Kancellári Hivatal és a Közlekedési, Fejlesztési és Technológiai Szövetségi Minisztérium is.

Az esemény rangjához méltóan európai szervezetek és intézmények vezetői és legismertebb szakemberei, az osztrák tudományos élet meghatározó egyéniségei, tanszékvezető egyetemi professzorok, vezető osztrák statisztikusok vállaltak előadást.

A rendezvény két jól elkülöníthető, de egymáshoz szorosan kapcsolódó részből állt. Az első rész a „Kihívások a hivatalos statisztikai szolgálattal szemben az egyesült Európába” címet viselte és házigazdája az Osztrák Statisztikai Hivatal (Österreichisches Statistisches Zentralamt), megújult nevén a Statistik Austria volt. A nyitó előadást, amit az intézmény két főigazgatója *Ewald Kutzenberger* és *Gabriela Petrovic* tartott, annak szentelték az előadók, hogy megindokolják, milyen szemléletbeli változás magyarázza és tölti meg tartalommal a névváltoztatást. A hivatal hatóságként is értelmezhető működése helyett a szolgáltatás jelleget, a társadalom szolgálatát kívánták jelezni és hangsúlyozni a névváltoztatással.

Az első napon a hivatalos statisztikai szolgálattal szembeni kihívások megfogalmazását az EUROSTAT, az Európa Tanács, az Európai Központi Bank és az OECD vezetői és szakemberei tették teljessé. A nap második felében az Internetes közvéleménykutatások, a regiszterek aktuális kérdései kerültek megvitatásra.

A konferencia második napján tematikus szekciókban folytatódott a munka. Az első szekció a vállalkozásokkal, a második a regionális gazdasággal, a harmadik a foglalkoztatás kérdéseivel és a munkaerővel foglalkozott. A rendezvény záró előadása a „Politikai monitoring Európában” címet viselte.

A konferencia második részében a házigazda szerepét a jubiláns Osztrák Statisztikai Társaság vette át. Ewald Kutzenberger előadásában azt fejtette ki, hogy a jövőben is meghatározó lesz az Osztrák Statisztikai Hivatal (Statistik Austria) és az Osztrák Statisztikai Társaság partnerkapcsolaton alapuló együttműködése. *Herbert Mang* az Osztrák Tudományos Akadémia nevében üdvözölte az ünnepelt szervezetet, míg *Alexander Pinwinkler* a salzburgi egyetem nevében ismertette az Osztrák Statisztikai Társaság 30 éves történetét.

A tudományos ülés a hivatalos statisztikai szolgálat minősítése jegyében folytatódott. *Wilhelm Adema*, az OECD szakembere a fogyasztás szintjének méréséről értekezett. *Alfred Franz*, a Statistik Austriától „Hivatalos statisztika globális nézőpontból” címmel tartott előadást. *Erich Wonka* szintén a Statistik Austria munkatársa a hivatalos nemzeti statisztikai adatszolgáltatás globalizációjáról adott elő. A nap záróelőadását *Andreas Quatemberger*, a linzi egyetem professzora tartotta hat európai ország hivatalos statisztikájának összehasonlításáról.

Október 18-án a téma a gazdaság és szociálpolitika volt. A moderátor szerepét *Margit Epler*, az Osztrák Munkavállalók Kamarájának elnöke vállalta el. A délutáni program *Peter Hackl*, a Bécsi Egyetem professzorának moderálásával „Tematikus statisztika és a nyilvánosság” címet kapta. Az öt elő

adás között a Magyar Statisztikai Társaság képviselőjében hangzott el *Herman Sándor* „Az alacsony aggregált statisztikai adatok elemzésének néhány kérdése” című előadása, így téve eleget az Osztrák Statisztikai Társaság felkérésének.

Az utolsó nap az új statisztikai módszerek bemutatása jegyében történt. Az előadások moderáto-

rai *Reinhard Viertl* és *Rudolf Dutter*, a Bécsi Műszaki Egyetem professzorai voltak. A rendezvényen osztrák, francia, angol, spanyol, olasz, svájci és magyar előadók vettek részt, a konferencia nyelve német, illetve angol volt.

*Herman Sándor*

## SZEMÉLYI HÍREK

**Felmentés-megbízás.** *Dr. Mellár Tamás*, a Központi Statisztikai Hivatal elnöke *Imre Jánosnak* a Műszaki és Ellátási főosztály vezetésére kapott megbízásáról szóló lemondását 2002. január 31-ével tudomásul vette, és 2002. március 14-ei hatállyal megbízta *Kreutzer Aurélt* a Műszaki és Ellátási főosztály vezetésével.

\*

*Helt Ferenc*, a Központi Statisztikai Hivatal elnökhelyettese *Koltai Tamástól* az Informatikai főosztály főosztályvezető-helyettesi feladatok ellátására adott megbízását 2002. január 31-ei hatállyal visszavonta.

*Dr. Soós Lőrinc*, a Központi Statisztikai Hivatal elnökhelyettese *dr. Horváth Józseftől* a sajtó és kommunikációs szaktanácsadói feladatkörre adott főosztályvezető-helyettesi megbízását 2002. február 28-ai hatállyal visszavonta.

*Dr. Bagó Eszter*, a Központi Statisztikai Hivatal elnökhelyettese *Bedekovics Istvánt* a Nemzeti Szám- és Számítógéptudományi Hivatal főosztályon 2002. február 1-jei hatállyal megbízta a főosztályvezető-helyettesi feladatok ellátásával.

**Kitüntetés.** *Dr. Mellár Tamás*, a Központi Statisztikai Hivatal elnöke *Varga E. Árpádnak*, a Kulturális Innovációs Alapítvány Könyvtára könyvtárvezetőjének, Erdély népességi viszonyainak részletes, településszintű megismertetésében végzett – külföldön is elismert – jelentős munkássága elismeréseképpen *Fényes Elek Emlékéremet* adományozott.

**Elnöki dicséret.** *Dr. Mellár Tamás*, a Központi Statisztikai Hivatal elnöke *Déményné Lehel Zsuzsannát*, a Tájékoztatási főosztály statisztikai tanácsadóját és *Holka Gyulát*, nyugdíjas főosztályvezetőt a „Századok statisztikája” c. kiadvány elkészítésében vállalt lelkiismeretes munkájuk elismeréseképpen; *Barabás Istvánnét*, a Szabolcs-Szatmár-Bereg Megyei Igazgatóság igazgatóhelyettesét, *dr. Bánki Katalint*, a Csongrád Megyei Igazgatóság osztályveze-

tőjét, *Hajdú Györgynét*, az Informatikai főosztály osztályvezetőjét, *Hauzer Györgynét*, a Veszprém Megyei Igazgatóság vezető főtanácsosát, *Horváth Zsoltnét*, a Fejér Megyei Igazgatóság igazgatóhelyettesét, *Kemény Endrénét*, a Bács-Kiskun Megyei Igazgatóság osztályvezetőjét, *Leányvári Lászlót*, a Borsod-Abaúj-Zemplén Megyei Igazgatóság igazgatóhelyettesét, *Szabó Ferencet*, a Mezőgazdasági főosztály tanácsosát a Szőlő- és gyümölcsös ültetvények összeírási munkáiban végzett eredményes munkájukért; *Kópházi Józsefet*, az Informatikai főosztály főosztályvezető-helyettesét, a statisztikai regiszter fejlesztésében 2001. évben végzett eredményes vezetői és szervezői munkája elismeréseként; *Kassainé Nagy Beátát*, az Informatikai főosztály főtanácsosát a Népszámlálási képviselői mintafeldolgozásában, valamint a teljes körű feldolgozásban végzett magas színvonalú munkájáért *elnöki dicséretben* részesítette.

**Jubileumi jutalmak.** Köszölgélati jogviszonyban töltött idejük alapján 2002. január–február–március hónapokban a Központi Statisztikai Hivatal következő dolgozói részesültek jubileumi jutalomban.

40 éves szolgálatért: *Farkas Gizella* (Nemzeti Számlák főosztály), *Kovács Ferencné* (Népesedés-, Szociális és Egészségügyi Statisztikai főosztály), *Könyves Gabriella* (Iparstatisztikai főosztály), *Szabó Ilona* (Informatikai főosztály).

35 éves szolgálatért: *Kalina Lászlóné* (Informatikai főosztály), *Kamarás Ferenc* (Népesedés-, Egészségügyi és Szociális Statisztikai főosztály), *Kiss Jánosné* (Informatikai főosztály), *Kolcza Klára* (Népszámlálási főosztály), *dr. Kulcsár Rózsa* (Mezőgazdasági főosztály), *Mogyorósiné dr. Halász Rózsa* (Elnökhelyettesi titkárság), *Széll Lajosné* (Pénzügyi főosztály).

30 éves szolgálatért: *dr. Bosznayné Gilicze Márta* (Informatikai főosztály), *Habáné Kamarás Zsuzsanna* (KSH Könyvtár és Dokumentációs

Szolgálat), *dr. Joubert Kálmán* (KSH Népeségutómányi Kutatóintézet), *Molnár Károlyné* (Pénzügyi főosztály), *dr. Rudas Jánosné* (Pénzügy-statisztikai főosztály), *dr. Vigh Judit* (Nemzeti Számlák főosztály), *Zsembrovzsky Irén* (Informatikai főosztály).

25 éves szolgálatért: *Baloghné Bartók Anna* (Nemzeti Számlák főosztály), *Csapó Éva* (Szolgáltatásstatisztikai főosztály), *Erdei Éva* (Pénzügyi főosztály), *Faragó Eszter* (Területi és Koordinációs főosztály), *Nagy Ferenc* (Műszaki és Ellátási főosztály), *dr. Szunyogh Zsuzsanna* (Életszínvonal- és Emberierőforrás-statisztikai főosztály).

#### SZERVEZETI HÍREK – KÖZLEMÉNYEK

**Kőrösy József-díj.** *Dr. Mellár Tamás*, a Központi Statisztikai Hivatal elnöke a Központi Statisztikai Hivatal területi szervezeti egységei munkájának szervezésében és irányításában, valamint a területi statisztika fejlesztésében több évtizeden keresztül kiemelkedően eredményes tevékenységet végzett köztisztviselők munkájának elismerésére *Kőrösy József-díjat* alapított. A díj adományozásáról a KSH elnöke dönt. A díj a KSH-val közszolgálati jogviszonyban álló, illetve a KSH-tól nyugdíjba vonult köztisztviselők részére adományozható. A díj oklevéllel és pénzjutalommal jár, amelynek nettó összege a mindenkor illetményalap háromszorosa. A díj odaítélésére irányuló javaslatokat a Területi és Koordinációs főosztályra kell eljuttatni. Az érdemes személyekre az e célra létrehozott bizottság minden év március 31-ig tesz javaslatot a KSH elnökének. Az elnök a díjat általában minden év április hónapban adja át.

**A Statistics Canada tervezési konferenciáját** 2002. január 14. és 15-én tartották Ottawában, melyen megfigyelőként *dr. Vukovich Gabriella*, a Központi Statisztikai Hivatal elnökhelyettese és *Laczkó Sándorné*, a Központi Statisztikai Hivatal főosztály-vezetője vettek részt.

A Statistics Canada éves tervezési konferenciájának célja, hogy a szakterületek vezetői bemutassák a stratégiai tartalék terhére a következő év folyamán tervezett feladatokat. A konferencia résztvevői ezeket a javaslatokat megvitatták, a hivatal vezető munkatársai pedig a terveket megismerve véleményezték azokat. A tervezési konferenciának nincs döntéshozó szerepe, főleg a későbbi döntéseket segíti elő, viszont van egy további és a tervezéssel egyenrangú fontosságú szerepe is. A konferencia ugyanis szeminárium jellegű és fórumot teremt arra, hogy minden szakterület beszámoljon a tervek mellett fontosabb eredményeiről és gondjairól. A hivatal vezető munkatársai így közvetlenül is megismerik a többi szakterületen folyó munkákat, az ott felmerülő nehézségeket és elért sikereket. Ezáltal a konferencia a szakmai kohézió erősítésének egyik eszköze.

Ott-tartózkodása során a magyar delegáció tagjai találkoztak *Richard Barnabéval*, a Kanadai Statisztikai Hivatal kommunikációs és koordinációs elnökhelyettesével, akivel a kanadai átvilágítást követő munka stratégiai kérdéseit tekintették át.

A kanadai hivatal mezőgazdasági statisztikusaival félnapos megbeszélésre került sor, mely lehetőséget adott kanadai mezőgazdasági statisztika rendszerének áttekintésére, különös tekintettel a kliens-éledegeltségi felmérésükre.

Ezt követően a KSH képviselői találkoztak a „Special Survey Division” részleg vezetőjével. Ez a részleg a külső megrendelők kérésére végrehajtott összeírások tervezésével, végrehajtásával foglalkozik.

A magyar vendégek megtekintették a kanadai hivatal oktatási-képzési központját, ahol a „Human Resources Development” részleg vezetője a hivatal oktatási és képzési tevékenységéről adott átfogó ismertetést. A vendégek a hivatal adat- és metaadatbázis-rendszeréről is tájékoztatást kaptak. A montreali területi igazgatóságon tett látogatás során a küldöttség tagjai a területi szervek működését, valamint a központ és a területi szervek közötti munkamegosztást tanulmányozták.

**A szociális statisztikai évkönyv 2000** első három fejezete a népesség demográfiai és foglalkoztatottsági jellemzőit mutatja be. A negyedik fejezet egyes veszélyeztetett társadalmi csoportok adatait teszi közzé. Az ötödik fejezet pedig az államháztartás pénzügyi adataitól kezdve tárja az olvasó elé a különböző pénzügyi ellátások adatait. Ezt követően a gyermekjóléti és -védelmi ellátást a szociális alap- és nappali ellátást, a rehabilitációs foglalkoztatást és a nonprofit szervezetek tevékenységének adatait tartalmazza az évkönyv, a szöveges részt angol nyelven is.

(Szociális statisztikai évkönyv, 2000. Központi Statisztikai Hivatal, Budapest, 2001. 235 old.)

**Egészségügyi statisztikai évkönyv 2000.** A kiadvány az összefoglaló adatokat követően közlést tesz a népesség főbb demográfiai jellemzőit, az egész-



ségügyi alapellátás, a járóbetegek szakellátása, a gondozóintézeti ellátás, valamint a sportegészségügy és a fogászat adatait. Részletes táblákat közöl a kötet a fekvőbeteg-gyógyintézeti, a laboratóriumi és a gyógyszerellátásról, a közegészségügyről és a járványügyről. Tájékoztatást nyújt a balesetekről, az öngyilkosságokról, az orvosszakértői tevékenységről, valamint az egészségügyi ellátás személyzetéről, az egészségügyi képzésről és továbbképzésről. Részletes adatok találhatóak az évkönyvben az egészségügyi biztosításról, az árak, keresetek és tárgyi eszközök alakulásáról, valamint nemzetközi összehasonlítható adatokról.

A kiadványt módszertani magyarázatok zárják.

(Egészségügyi statisztikai évkönyv 2000. Központi Statisztikai Hivatal. Budapest. 2001. 448. old.)

**A munkaerő-felmérés regionális idősorai 1992–2000 (Adattár).** Az adatsorok 1992–2000 között mutatják be a foglalkoztatás, a munkanélküliség és a gazdasági aktivitás legfontosabb jellemzőit. Újdonság, hogy regionális adatokat is közlése az adatgyűjtemény, mégpedig a munkaerő-felmérés megszokott publikációs rendje szerinti részletesebb bontású (kor, nem, iskolai végzettség, gazdasági ágak, foglalkozási főcsoportok) táblák egységes szerkezetű közzétételével. A kötet nem tartalmazza a munkaerő-piaci szempontból foglalkoztatottnak tekintendő sorkatonák adatait. A táblák és ábrák szövege angol nyelven is olvasható.

(A munkaerő-felmérés regionális idősorai 1992–2000. (Adattár) Központi Statisztikai Hivatal. Budapest. 2001. 159 old.)

**A halandóság földrajzi különbségei Magyarországon, 2000** című kötet a halandósági táblák adatait tartalmazza a 2000. évre. A koréves halandóság táblák adatait az össznépesség, a férfi-, illetve a női népesség általános halandóságának tisztított mutatóit tartalmazza. A rövidített halandósági táblák a férfi-, illetve női népesség területi halandósági különbségeiről tájékoztatnak. A kötetet angol nyelvű szöveges rész és módszertani megjegyzések egészítik ki.

(A halandóság földrajzi különbségei Magyarországon, 2000. Központi Statisztikai Hivatal. Budapest. 2001. 119 old.)

**Megjelent a Demográfia** 2001. évi 3–4. száma. A Tanulmányok rovatban jelent meg *Karl Schwarz*: 2000. évi jelentés a németországi demográfiai helyzetéről; *Klinger András*: Halandósági különbségek Magyarországon iskolai végzettség szerint; *S. Molnár Edit*: A közvélemény gyermekszám-preferenciái

és *Frey Mária*: Egyensúlyt teremteni a fizetett munka és a családi élet között, című tanulmánya. A Közlemények rovatban *Spéder Zsolt*: Életünk fordulópontjai. Az NKI Társadalmi és Demográfiai Panel-felvételének (TDPÁ) kutatási koncepciója és kérdőívének vázlatos ismertetése; *Balázs Lajos*: Néprajzi tűnődések a népesedésről (II.), *Pozsgai Péter*: Család és háztartás. Magyarországi történeti demográfiai források fogalomhasználata a 16–19. században című írása olvasható. A Figyelő rovat tájékoztatást nyújt *dr. Kovácsics József*: Szentgotthárd és környéke. Szentgotthárd környéki, dél-burgenlandi, őrségi és vendvidéki falvak és nemzetiségek (1183–1995) című Településtörténeti lexikonáról, a megyei történeti statisztikai helységnévtárak megjelenése kapcsán szervezett konferenciáról, valamint egy hozzászólást közöl a *Demográfia* 2000. évi 4. számában megjelent Azilum c. interjúhoz. A folyóirat e számát is folyóiratszeme és szakirodalmi ismertetések zárják.

**A környezetstatisztikai adatok 2000** című kötet részletes adatokat tartalmaz a természeti erőforrások készleteiről és felhasználásukról, a környezet terhelő kibocsátásokról, a környezet állapotáról és minőségéről, a természetvédelemről, a környezetvédelmi ráfordításokról. Közlése a népességről szóló alapvető háttéradatokat és a környezetre ható gazdasági tényezők alakulásának jellemzőit, valamint Magyarország néhány földrajzi adatát. A kötetet Magyarország meteorológiai adatai teszik teljessé. A táblák és ábrák szövege angol nyelven is megjelent.

(Környezetstatisztikai adatok 2000. Központi Statisztikai Hivatal. Budapest. 2002. 198 old.)

**Az Általános Mezőgazdasági Összeírás 2000** c. kiadvány az egyéni gazdaságok munkaerő-felhasználásának eredményeit dolgozza fel. Az elemzés öt fő témaköre: 1. A mezőgazdaság szerepe a nemzetgazdaságban; 2. A lakosság mezőgazdasági tevékenysége; 3. A munkát végzők főbb adatai; 4. Az egyéni gazdaságok munkaerejének jellemzői; 5. A munkaerő-felhasználás és a földterület kapcsolata. Az elemzést bőveges táblaanyag és módszertani megjegyzések egészítik ki.

(Az egyéni gazdaságok munkaerő-felhasználása 2000. Központi Statisztikai Hivatal. Budapest. 2001. 129 old.)

**Migráció és statisztika** címmel *Illés Sándor* és *Lukács Éva* szerkesztésében tanulmánykötet jelent meg **A személyek szabad áramlásának statisztikai szempontú vizsgálata** című projekt eredményeiről. A projekt a Szociális és Családügyi Minisztérium, valamint a Központi Statisztikai Hivatal együttműködése keretében valósult meg. Az együttműködés ku-

tatási részében a következő nyolc téma kidolgozása valósult meg: 1. A külföldön tartózkodó magyar állampolgárok felmérése a kilencvenes években, különös tekintettel az Európai Unió tagállamaiban tartózkodókra, továbbá a visszavándorló (hazatérő) magyar állampolgárok vizsgálata; 2. Kivándorló magyar állampolgárok vizsgálata a főbb európai országokban nevezetesen, Németországban és Ausztriában; 3. Az Európai Unió tagállamaiból bevándorló és Magyarországon munkát vállaló uniós polgárok; 4. A gyermek vándorlása – családgyesítés; 5. A

külföldön tanuló magyar diákok és a Magyarországon tanuló külföldi diákok; 6. A külföldön munkát vállaló magyar munkaerő, különös tekintettel Németországra és Ausztriára; 7. A magyar nyugdíjasok nemzetközi vándorlása; 8. A kapcsolat a hivatalos EU-vélemények és aközött, hogy az adott tagállamok közvéleménye hogyan ítéli meg a csatlakozni kívánókat.

(Migráció és statisztika. Szerk.: *Illés Sándor, Lukács Éva*. Népeségtudományi Kutatóintézet. Kutatási jelentések 71. Budapest, 2002. 246 old.)

# STATISZTIKAI IRODALMI FIGYELŐ

## KÜLFÖLDI STATISZTIKAI IRODALOM

### A STATISZTIKA ÁLTALÁNOS ELMÉLETE ÉS MÓDSZERTANA

BLACKWELL, L.:

#### A NŐK MUNKAVÉGZÉSE AZ EGYESÜLT KIRÁLYSÁGBAN ÉS A FOGLALKOZÁSOK ÚJRAOSZTÁLYOZÁSA

(Women's work in UK official statistics and the 1980 reclassification of occupations.) – *Journal of the Royal Statistical Society*, 2001. 2. sz. 307–325. p.

A társadalom foglalkozási struktúrájának hosszú távú vizsgálatát gyakran igen megnehezíti, hogy a foglalkozások osztályozásának módosításakor nem dokumentálják kellőképpen sem a változások tartalmát, sem a társadalmilag indokolt finomítások lényegi mérveit. A kutató számára így nehezen megoldható feladat a tényleges szerkezeti eltolódások megkülönböztetése azoktól a „művi” hatásoktól, amelyek pusztán a megváltozott csoportosítások következményei.

A tanulmány az Egyesült Királyság Statisztikai Hivatalának az 1971 és 1991 közötti időszakot átfogó longitudinális kutatási adataira támaszkodva vizsgálja, hogy a foglalkozások osztályozásának 1980-ban végrehajtott módosításai hogyan befolyásolták a nemek szerinti foglalkozási arányokat. A longitudinális vizsgálat Anglia és Wales népességének 1 százalékára vonatkozott, azaz az 1971, 1981, illetve 1991. évi censzók alkalmával egyaránt mintegy 500 ezer főre terjedt ki. A személyi adatokhoz kapcsolódóan bizonyos kiegészítő információkkal is rendelkeztek (például az életkor, a házastárs halála, az élve, illetve halva született gyermekek száma).

A foglalkozási struktúra idősorainak vizsgálatánál a fő probléma az, hogy a korábbi időszakokra vonatkozó adatok általában kevésbé részletezettek. Az 1971. évi census például még a foglalkozások 1970. évi osz-

tályozásán alapult, amely 223 foglalkozási alcsoportot különböztetett meg. Az 1981. évi censznál már a foglalkozások 1980. évi osztályozását használták, s ez alapján eredetileg 351 foglalkozási alcsoportot vettek figyelembe, majd számukat 549-re bővítették. Az ilyen jellegű összehasonlíthatósági problémák kezelésére a nemzetközi szakirodalomban háromféle stratégiai javaslat található. Az első eljárásnál minden érintett időszakban az akkor érvényes osztályozást használják (lehetőség szerint utalva a bekövetkezett módosításokra.). A második megoldás, hogy a foglalkozásoknak csak azt a szűkebb körét (az ún. magot) veszik figyelembe, amely közös valamennyi osztályozási rendszerben. Ilyenkor azonban a módosítások által érintett foglalkozások figyelmen kívül hagyása hátrányosan befolyásolja az eredmények reprezentativitását. A harmadik lehetőség, hogy – a részletesebb csoportosítások adatait megfelelően összevonva – a korábbi, kevésbé tagolt foglalkozási adatokat ezekből is előállítsák. Ekkor viszont a munkaerő összetételének éppen azok a változásai nem jutnak kifejezésre, amelyek a munkavégzés aktuális differenciálódásával kapcsolatosak.

Szem előtt tartva, hogy a nemek szerinti arányok módosulásának tisztázása az elsődleges célok közé tartozik, a longitudinális vizsgálat során az egyéni foglalkozásváltozás következő főbb típusait különböztették meg:

– A kutatókat leginkább érdeklő *első típust* a foglalkozás valódi változása jelentette (amikor például egy női munkaerő az 1971-ben még tipikusan „női” foglalkozását 1981-re más, jellegzetesen „férfi” szakmára cserélte fel.)

– A *második típus* esetén az egyén továbbra is ugyanazt a foglalkozást folytatta, de a két census közötti időszak-

*Megjegyzés.* A *Statisztikai Irodalmi Figyelő* rovatot a Központi Statisztikai Hivatal Könyvtár és Dokumentációs Szolgálat állítja össze. A rovat minden hónapban *Külföldi Statisztikai Irodalom* fejezetet (külföldi statisztikai és demográfiai könyvek és cikkek ismertetését *Rettich Béla* szerkesztésében), páratlan hónapban *Bibliográfiát* (a könyveket az MSZ 3423/2–84, az időszaki kiadványokat az MSZ 3424/2–82 szabvány szerinti feldolgozásban), páros hónapokban *Külföldi folyóiratszémlelét* tartalmaz.

ban végbement nemek szerinti összetétel-változás következtében a foglalkozás jellege változott meg: például a korábbi férfi túlsúllyal szemben a szakma „elnőiesedett”.

– A *harmadik típust* az az eset jelentette, amikor ugyanazt a foglalkozást a későbbi osztályozási rendszer máshova sorolta.

A tanulmány elsősorban ez utóbbi hatásának a számszerűsítésére törekedett, abból a célból, hogy az első és második típus szerepét torzítás nélkül, „tisztán” lehessen elemezni.

A foglalkozások egymást követő osztályozási rendszereinek összehasonlítása érdekében az Egyesült Királyság Statisztikai Hivatala az 1971. évi cenusból nyert egyszázalékos minta adatait (amelyek a teljes, illetve rész munkaidőben foglalkoztatott férfiakra és nőkre vonatkoztak), a foglalkozások 1980. évi osztályozási rendszerének megfelelően, másodlagosan is feldolgoztatta. A foglalkozások nemek szerinti koncentrátságának vizsgálatához a szerző az ún. Warwick konverziós program segítségével (amelyet *Peter Elias* és társai, a Warwick Egyetem Munkaügyi Kutató Intézetének munkatársai fejlesztettek ki) oldotta meg az 1981. évi minta visszamenőleges átdolgozását az 1970. évi osztályozási rendszer szerint. Az 1991. évi adatok feldolgozása eleve párhuzamos (az 1980. évi csoportosításokat is figyelembe véve) kódolás alapján történt.

A szerző tanulmányában a szigetországban 1811 és 1831. között végrehajtott censzusoktól kezdődően tekinti át a női foglalkoztatottság főbb jellemzőit, és ismerteti egyéni munkájuk számbavételének – jórészt az összeírási előírásokból adódó – kezdeti korlátait. A XIX. században egyértelműen a háztartási alkalmazottként végzett munka jelentette a legáltalánosabb női foglalkozást s még a XX. század első felében is főleg az egyedül álló nők körére koncentráldott a női munkavállalás. (A férjzett nők közszolgálati alkalmazásának tilalma például egészen 1946-ig érvényben volt.) A jellegzetesen férfi foglalkozások viszonylag részletes osztályozása kezdettől fogva a férfiak magasabb szakképzettsége iránti igény elismerését fejezte ki s e hagyományos megkülönböztetés jórészt még ma is észlelhető. A tipikusan női foglalkozások néhány foglalkozási alcsoportba tömörítése gyakorlatilag szintén a nők társadalmi alulértékelését tükrözi, amely legfeljebb betanított munkának minősíti a főleg nők által betöltött munkaköröket.

A foglalkozások 1970., illetve 1980. évi osztályozási rendszerére egyaránt jellemző a jellegzetesen férfi, illetve női foglalkoztatás polarizálódása. A férfiak több mint fele olyan foglalkozást folytat, ahol a férfiak részesedése meghaladja a 90 százalékot, míg a női munkavállalók egynegyedének van olyan fog-

lalkozása, ahol a nők aránya legalább 90 százalék. Különösen erős a nők tömörülése az irodai foglalkozások és pénztárosok, valamint az alap- és középfokú iskolák tanítói és tanárai között, ahol részesedésük 60-70 százalék között mozog.

A Warwick konverziós program alkalmazása lehetőségét nyújtott a foglalkozások 1970. és 1980. évi osztályozásainak tételes megfeleltetése révén olyan – nemek szerint elkülönülő –  $10 \times 10$ -es matrixok összeállítására, amelyek sorai az 1970. évi, oszlopai pedig az 1980. évi arányok szerint ábrázolják a foglalkozási csoportok megoszlásait. A főtló adatai azokra a munkavállalókra vonatkoznak, akik besorolásában az osztályozás módosítása nem idézett elő változást. A főtlón kívül eső adatok az átsorolások aránymódosító hatásait fejezik ki, azokra az esetekre vonatkozóan, amikor maga a munkavállaló nem változtatott foglalkozást, de az osztályozási rendszer átalakítása módosította hovatartozását.

A férfiak 77 százalékát nem érintették az osztályozási rendszeren 1980-ban végrehajtott módosítások, s a fennmaradó 23 százalék esetében is előfordult, hogy a más kategóriába kerülés alapvetően a nők arányának változásával függött össze. Így például 1971-ben a műszaki rajzolók között a nők 11 százalékos aránnyal szerepeltek. Az 1980. évi osztályozási rendszer szerint viszont részesedésük 3 százalékra csökkent, mivel többségüket átsorolták az előrajzolókat, irodai kisegítő rajzolókat magába foglaló foglalkozási csoportba, amelyre korábban is a női túlsúly volt jellemző. A nők foglalkozás szerinti koncentráldását bemutató mátrix tanúsága szerint a női munkavállalóknak csak 22,6 százalékát nem érintette a foglalkozások osztályozásának 1980. évben hatályba léptetett felülvizsgálata. Lényeges átrendeződéssel járt például a takarító foglalkozások meghatározásának módosítása. Az 1970. évi változat szerint e kategóriában 86 százalékos női túlsúly volt jellemző, mivel a lakás- és irodai takarítónők, ablaktisztítók és kéményseprők tartoztak ide. 1980-tól idesorolták az utcaseprőket, valamint az üzemi takarítókat is, ami számottevően csökkentette a nők arányát és széles körben átsorolásokat tett szükségessé, holott az érintettek által végzett tevékenység teljesen azonos maradt a korábbival.

A longitudinális vizsgálatoknál jól használhatónak bizonyult olyan nemek szerint elkülönülő, egyszerűsített,  $3 \times 3$ -as matrixok összeállítása is, amelyek sorai az 1970. évi, oszlopai pedig az 1980. évi arányoknak megfelelően ábrázolják azokat a foglalkozás csoportokat, amelyekre vagy a férfiak, vagy a női munkavállalók túlsúlya (azaz 70 százaléknál magasabb aránya) a jellemző, illetve amelyek ebből a szempontból vegyes tartalmú csoportnak tekinthe-

tők. Ez az elhatárolás jó szolgálatot tehet olyan vizsgálatoknál is, amelyek például az ipari erőviszonyokat, a béregyeztetéseket, munkafeltételek alkupozícióit kívánják jellemezni. Ugyanakkor a foglalkozási osztályozások módosításának hatásait ezek a nagy, átfogó foglalkozási kategóriák csak tompítottan érzékelhetik. A főatlóban szereplő (vagyis a besorolási változások által nem érintett) adatok összege ugyanis a férfiak mátrixában 92, a nők mátrixában 90 százalékos arányt jelez.

A longitudinális vizsgálatok keretében az előzőkhöz hasonló számítások készültek azoknak az 1931 és 1951 között született munkavállalóknak a kohorszára vonatkozóan is, akikre mindkét census kiterjedt. Életkoruk 1971-ben a 20-39 éves, 1981-ben a 30-49 éves intervallumba tartozott. A kutatók az egyszázalékos minta és a kohorsz foglalkozási struktúrájának valószínű eltérését elhanyagolható nagyságrendűnek vélelmezték.

A foglalkozásistruktúra-változás és a kódolási hibák hatásának kiszűrése után kapott eredmények arra mutatnak, hogy mind a nők, mind a férfiak körében növekszik azoknak az aránya, akik saját kezdeményezésből nem cserélik foglalkozásukat. Összességében azonban a nők gyakrabban változtatnak: a férfiak háromnegyed részének volt ugyanaz a foglalkozása mindkét census alkalmával, a nőknél viszont ez az arány csak kétharmadot képviselt. Hasonló tendenciát jelez, hogy míg a férfiak kétharmada maradt meg a jellegzetesen férfi foglalkozások körében, addig a nők kevesebb mint egyharmadának volt tipikusan női foglalkozása mindkét census időszakában.

(Ism.: *Tűű Lászlóné*)

OMAR, E.:

A STATISZTIKA IRÁNTI NÖVEKVŐ IGÉNY  
AZ EMBERI JOGOK TERÜLETÉN

(Development and human rights: The growing demand for statistics from the international community.) - *Statistical Journal of the United Nations ECE*. 2001. 2-3. sz. 141-153. p.

Az emberi jogok védelméért való összefogás újabban nem annyira jogi, mint inkább gazdasági, szociális és kulturális kérdésekre összpontosít. Más megfogalmazásban az emberi jogok és a társadalmi-gazdasági fejlődés kapcsolatának vizsgálata került előtérbe. Az emberi jogok két dimenzióban érvényesülnek: a félelem nélküli életben és a szükségétől mentes életben. Az ezekkel kapcsolatos megvalósítási mechanizmusokat a Polgári és Politikai Jogok Nemzetközi

Konvenciója, illetve a Gazdasági, Szociális és Kulturális Jogok Nemzetközi Konvenciója rögzíti.

A demokráciának biztosítani kell a szabad választást, az egyesülési szabadságot, a független igazságszolgáltatást, a kormány számonkérhetőségét, az információkhoz való szabad hozzáférést, a hatékony államigazgatást, a kormányzat és a civil szervezetek közötti együttműködést. A statisztikusok egy része éppen az ilyen jellegű adatok hiányában kerül az ilyen széles körű értelmezést, amikor mérésre, mutatók számítására vállalkozik.

Az emberi jogok védelmének eszközei között a statisztika mind ez ideig nem kapott hangsúlyos szerepet. Ennek az lehet az egyik oka, hogy a statisztikai mutatószámok nehezen értelmezhetők a média és a nagyközönség számára. Ugyanakkor aligha képzelhető el az emberi jogok érvényesülésének figyelemmel kísérése és a helyzet javítását célzó politika kidolgozása az e célt szolgáló statisztikai rendszer nélkül. A statisztika több területen is hozzájárulhat a helyzet feltáráshoz, így például a népesedés, a környezet, a foglalkoztatottság, a háztartás és család, a települési adatok, az egészségügy, az oktatás, a kultúra vagy az érdekérvényesítés módjai tekintetében.

Ami a szükséges intézményi rendszert és annak működését illeti, szakmailag jól felkészült statisztikai hivatalokra van szükség, ahol a különböző felek által elfogadott szabványos módszereket alkalmaznak a kérdések vizsgálatára. A munkának természetesen költségigénye is van. Valamely kérdést vizsgáló nemzetközi felmérésre gyakran a külügyminisztériumok válaszolnak, mivel az emberi jogok érvényesülését politikai kérdésnek tekintik. A statisztikai háttér megléte ilyenkor másodlagos szempont.

Az emberi jogok megsértésének bemutatására egyszerű módszerek is alkalmasak, mint például fénykép, interjú, beszámoló. A helyzet javítását célzó politika kialakításához, az elérendő célok pontos megfogalmazásához és az elért eredmények figyelemmel kíséréséhez azonban ennél többre van szükség: egységes módszerekkel gyűjtött adatokat szolgáló statisztikai rendszerre.

A szükséges adatokat gyűjtő és feldolgozó statisztikai szervezetnek jól kell kommunikálni a kormányzat, a gazdasági döntéshozók, a civil szervezetek és a lakosság felé. Rá kell mutatni az adatok helyes értelmezésére, részben azért, mert újszerűek lehetnek, részben pedig azért, mert szerteágazó összefüggéseik vannak, amelyek között a jogrendszer és a jogalkalmazás csak az egyik terület. A politikai felsővezetésnek biztosítani kell, hogy a statisztikai szervezetet nem éri retorzió, amennyiben a politikai vezetők számára zavaró adatokat tesz közzé. A nemzeti

statisztikai hivatalok ilyen vonatkozásban számíthatnak a nemzetközi szervezetek támogatására.

Az emberjogi nemzetközi szerződések aláírójának időről időre jelentést kell készíteniük az emberi jogok érvényesüléséről a kérdéses országban, amire a szerződésben kötelezettséget vállaltak. A jelentések formája, éppen a nemzetközi összehasonlíthatóság érdekében kötött és rendszerint kérdésekre kell válaszolni. A kérdések értelmezését a szerződés megfelelő részeire való hivatkozással teszik félreérthetlenné. Az eredmények összesítése a nagyközönség számára elérhető formában jelenik meg.

A hivatkozott jelentésekhez a következő területekről várnak statisztikai adatokat: foglalkoztatottság, munkanélküliség országos szinten és egyes csoportok (nőket, fiatalok, idősek, mozgássérültek) vonatkozásában; életszínvonal és életminőség; részletes adatok a lakásviszonyokról; csecsemőhalandóság (nemek szerint, város–vidék részletezésben, társadalmi–gazdasági és etnikai csoportok szerint, területi bontásban); ivóvíz-ellátottság (város–vidék részletezésben); szemét- és hulladékkezelés (város–vidék részletezésben); iskolázottság, felnőttoktatás, továbbképzés (lehetőleg nem, vallás stb. szerinti bontásban), várható élettartam (város–vidék, társadalmi–gazdasági csoportok és nemek szerinti részletezésben).

Ezen adatok jó része a kialakult statisztikai rendszerekben rendelkezésre áll, legfeljebb a részle-

tezéseknél kell tekintettel lenni az emberjogi vonatkozásokra. Kevésbé fejlett statisztikai rendszerrel bíró országoknak nemzetközi szakértői segítségre van szükségük. A statisztikai mutatóknak végül is négy kérdéskörre kell rávilágítani: 1. az emberi jogok körére, 2. az állam által vállalt kötelezettségekre, 3. az elért eredményekre és 4. az időbeli fejlődésre. Mindezek lehetnek mennyiségi (számszerű) és minőségi (szövegesen értékelhető) mutatók.

Az emberjogi szerződésekben előírt jelentésekhez mintegy 130 statisztikai mutató számíthatóságát több mint 20 országban tesztelik. Úgy tűnik, 50 mutatószám általában alkalmazhatónak látszik. Ugyanakkor a tervezett listában nem szereplő további 50 mutató is szóba jöhet az egyes országok saját viszonyainak felmérésénél. (A munkáról lásd [www.unep.ch/earthw/indstat.htm](http://www.unep.ch/earthw/indstat.htm).)

A szerző javasolja, hogy legyen egy minimális listája az emberi jogokra vonatkozó statisztikai mutatószámoknak, melyet a fejleményeknek megfelelően ötvenként felülvizsgálnának. Megfelelő indexszám kidolgozására is szükség lenne. Az emberjogi szervezeteknek az eddiginél aktívabb szerepet kell vállalniuk a világkonferenciákon, az ENSZ statisztikai szervezeteiben, az ENSZ által közreadott Human Development Report formálásában és általában a statisztikusokkal való együttműködésben.

(Ism.: Szász Kálmán)

## GAZDASÁGSTATISZTIKA

AYHAN, H. ÖZTAS:

### A NEMEK ASZIMMETRIJÁNAK CSÖKKENTÉSE A MEZŐGAZDASÁGI ÖSSZEÍRÁSOKBAN.

(Statistics by Gender: Measures to Reduce Gender Bias in Agricultural Surveys.) – *International Statistical Review*. 2001. 3. sz. 447–460. p.

A nemek szerint csoportosított statisztikai adatok a politikai döntéshozók érdeklődésének előterében állnak az utóbbi években. Ezért az összeírók arra törekednek, hogy az adatok ilyen bontásban is rendelkezésre álljanak. A tanulmány olyan javaslatokat tartalmaz, amelyek a fejlődő országokban a mezőgazdasági összeírásoknál csökkenthetik a nemek közötti aszimmetriát.

A nemek szerinti különbségtétel a nők és a férfiak helyzetére vonatkozó társadalmi konstrukció. A tévesen gyakran szinonimákként használt *sex* és *gender* kifejezések között különbséget kell tenni. A *sex* egyszerűen a nők és a férfiak közötti biológiai

eltérésekre utal. A nemi jellemzők univerzálisak és megváltoztathatatlanok, míg a *gender* társadalmi konstrukció és azoknak a különbségeknek a kodifikációja, amelyek a nemek, valamint a nők és férfiak közötti társadalmi kapcsolatok között fennállnak.

Első ízben 1975-ben az Egyesült Nemzetek Szervezete fejezte ki azt, hogy szükség van a nemek szerint bontott adatokra. Azóta az ENSZ különféle határozatai és ajánlásai bizonyos javulást eredményeztek a nemek szerint bontott adatok elérhetőségében. Mindazonáltal a nőkre vonatkozó adatok köre még igen szegényes.

A szóban forgó nemzetközi erőfeszítések részeként a FAO (Food and Agriculture Organization) keretében kutatások kezdődtek a nemek szerint csoportosított adatok szélesebb körű előállítására és felhasználására. A kutatási jelentések a Közel-Kelet számos országának a tapasztalatain alapulnak. Továbbá közzétették a vonatkozó latin-amerikai és karibi tapasztalatokat is a mezőgazdasági összeírások terén.

Az ENSZ, az Európai Gazdasági Bizottság és a Svéd Statisztikai Hivatal kiadványt tett közzé az európai és az észak-amerikai nőkkel és férfiakkal kapcsolatos válogatott mutatószámokról. Az ENSZ továbbra is érdeklődést mutat a kérdés iránt és az erre vonatkozó 1995. évi jelentésében különös figyelmet kapott a nemek szerinti egyenlőtlenség mérése. Ezen túlmenően a világ női lakosságával foglalkozó speciális kiadványban összefoglalták azokat a trendeket és országstatisztikákat, amelyek a női munkaerőforrásban játszott szerepével kapcsolatosak.

A nemek szerinti adatok elérhetőségével kapcsolatban megállapítható, hogy a korai fejlődési periódusokban a nők termeléshez való hozzájárulása nem volt mérhető a megfelelő információk hiánya miatt. Következésképpen a nemek szerint csoportosított adatok és statisztikák gyűjtésével és bemutatásával kapcsolatos irodalom nagy része aszimmetriára mutatott rá a nők adatainak kihagyása miatt.

A nemek szerint csoportosított statisztika előállításának célja a mindkét nemmel és nem csak a nőikkel kapcsolatos statisztika tökéletesítése.

A mezőgazdasági összeírások nemek szerint felbontott adatai vonatkozásában megfigyelhető problémák: *a)* A nemzetközi szervezetek adatgyűjtési programjaikban, ajánlásaikban és évkönyveikben nem eléggé hangsúlyozzák a nemek szempontjából érzékeny kérdéseket. *b)* Néhány fejlődő országban, ha a meglévő jogi, kulturális fogalmak nem egyeznek meg az elfogadott definíciókkal, ez statisztikai szempontból problémát okozhat. *c)* A legtöbb fejlődő országban a hivatalos statisztikai adatgyűjtés során nem fordítanak elég figyelmet a nemek szempontjából érzékeny kérdésekre. *d)* A nemekkel kapcsolatos aszimmetria csökkentése érdekében tett múltbéli erőfeszítések nem elégségesek.

A vázolt problémákkal kapcsolatban a szerző a fejlődő országok vonatkozásában részletesen megvizsgálja a mezőgazdasági összeírások nemekkel kapcsolatos érzékeny kérdéseit és az aszimmetria csökkentésének lehetőségeit. Ennek során rámutat a nemzetközi szervezetek módszertani dokumentációjának és ajánlásainak hiányára, majd ismerteti a fogalmakhoz, definíciókhoz, adatgyűjtéshez és táblakészítéshez kapcsolódó problémákat. Végül javaslatokat tesz a nemek aszimmetriájának mérséklésére és felszámolására az összeírás egyes szakaszaiban. Megállapítható, hogy a legtöbb múltbéli mezőgazdasági statisztikai összeírásnál a fejlődő országokban a nemekkel kapcsolatos kérdéseket, valamint az adatgyűjtési és a tájékoztatói fázisokkal való összefüggéseiket nem vették megfelelő mértékben figyelembe.

A mezőgazdasági összeírásoknál használt fogalmaknak és definícióknak semlegeseknek kell lenniük a

nemek vonatkozásában. A földtulajdonos személyének, a háztartás fejének egyértelműen férfinak vagy nőnek kell lennie függetlenül a kortól és a jövedelemtől. A nemzetközi ajánlások szerint a dolgozó személy definíciójának meg kell felelnie az ILO jelenlegi útmutatásainak. Különösen a segítő családtag fogalma jelent problémát a női munkaerő esetében, s erre a jövőbeli mezőgazdasági összeírásoknál különös figyelmet kell fordítani. E fogalom helytelen használata ugyanis nemcsak nemek szerinti aszimmetriát eredményez, hanem torzíja az adatokat is oly módon, hogy túlbecsli a nem fizetett női munkaerőt, a lehetséges női birtokok számát pedig alulbecsli.

A mezőgazdasági birtok működhet elkülönülő földtulajdonnal vagy anélkül, és a minimális földméret regisztrálandó a mezőgazdasági összeírások során. Fontos a nők mezőgazdasági termeléshez való hozzájárulásának kimutatása, mert a női hozzájárulás nagyobb lehet a föld nélküli birtokon és az olyan birtokokon, amelyek a minimális szint alatt vannak.

A statisztikai szakképzés hatékonyságát fokozni kell a nemek szerinti aszimmetria csökkentése céljából oly módon, hogy nagyobb figyelmet kell fordítani az olyan definíciókra, mint a háztartás feje, a birtokos, a birtok és a mezőgazdasági dolgozó. A mezőgazdasági összeírások minden szakaszában ösztönözni kell a női személyzet részvételét számlálóbiztosként, ellenőrként, kódolóként.

A táblák nemek szerinti bontásának rendelkezésre kell állnia az alapvető mezőgazdasági változókra a felhasználói igények szerint.

Mindazonáltal a nemek szerinti statisztikai kérdésekkel kapcsolatban a mezőgazdasági összeírások során a fejlődő országokban még további módszertani kutatásra van szükség.

(Ism.: Balogh András)

SÖNDERMANN, M.:

#### A KULTURÁLIS ÁGAZATOK ÉS A FOGLALKOZTATÁS EURÓPÁBAN

(Cultural industries and employment in Europe.) –  
*Cities and Regions*, 2001. 2. sz. 11–22. p.

Úgy tűnik, hogy a kultúra és a média napjainkban kulcsgázatokká váltak és egyre több embernek adnak munkát a jövőben. A művészet, a formatervezés, a multimédia, a zene és a digitalizáció, az irodalom és az internet divatszavak ebben a szöveggörnyezetben.

A kulturális ágazatok és ezek munkaerőpiaca azonban nem tartoznak a legfontosabbak közé sem a

politikai dokumentumokban, sem a stratégiai programokban. Ez azért is meglepő, mert a kultúrának és a médiának jelentős és növekvő fontosságú gazdasági és munkaerő-piaci szerepe van. Az Európai Unió Társadalmi Ügyek Főigazgatóságának 2001-ben publikált adatai szerint a kulturális szektor foglalkoztatási és gazdasági növekedése meghaladta az átlagot. Több mint 7 millió ember dolgozik az EU magán- és közösségi alapítású kulturális szektorában. A főigazgatóság közgazdászainak és munkaerő-piaci kutatóinak becslése szerint, a kisvállalkozásokat is beleértve a foglalkoztatottak teljes létszáma mintegy 12,4 millió fő.

Ha a kulturális szektor a médiával és a telekommunikációval ilyen figyelemre méltó eredményeket képes elérni, akkor úgy tűnik, nincs szükség speciális kulturális és gazdasági politikára. A kulturális ágazat alkotóművészei és más foglalkoztatottjai képesek megélhetésüket és munkájukat biztosítani e fellendülő szektorban. A növekedés hosszú távon is fenntartható, mivel a közgazdászok becslései szerint további 9,6 millió foglalkoztatott vonható be a kulturális szektorba.

A gazdaságban és a foglalkoztatottságban játszott figyelemre méltó szerepe ellenére a kulturális szektor mind a mai napig nem volt képes a nemzeti és az EU-beli döntéshozók figyelmének felkeltésére. Jól ismert tény, hogy a kulturális szektor igen összetett. Mégis van néhány eszköz a kultúra és a művészetek támogatására. A művészek, a kultúra döntéshozói és szakemberei közül számosan úgy gondolják, hogy a művészetek és a kultúra helyzete nem igényel empirikus elemzést. A művészek viszont nem akarnak művészeti munkaköröket létrehozni, de a művészetükből akarnak megélni. A művészeti stúdiók, a galériák, a művészeti egyesületek és a múzeumok, a zenekarok, a hangstúdiók, a zeneiskolák, az operaházak, az előadóművészek, a mozik, a tévé- és médiatársaságok már kifejlesztették átfogó hálózatukat Európa minden régiójában.

Mindenki, aki ebben a széles értelemben vett kulturális piacon dolgozik már hozzászólt a kultúra és a művészetek kettős jellemzőjéhez. A művészeti munkák és a kulturális szolgáltatások ugyanis kereskedelmi áruként is meghatározhatók. A Kereskedelmi Világszervezetben (World Trade Organization) vita folyik arról, hogy vajon a művészeti alkotásokkal, a kulturális termékekkel és szolgáltatásokkal ugyanúgy kell-e foglalkozni, mint a kereskedelmi árukkal, vagy a kulturális szektort speciális védelemben kell-e részesíteni. Ha kereskedelmi egyezményekkel megtiltanák a kulturális szolgáltatások támogatását, a közalapítású kultúra a legteljesebb veszélynek lenne kitéve.

A legtöbb európai országban a művészeteknek és a kultúrának három forrás képezi az alapját: az állam, a szabadpiac és a magán pénzalapok (szponzorálás, jótékonyosság). Ez tehát egy háromszektorú megközelítésnek felel meg, amelyet az elmúlt évek alatt fejlesztettek ki. A kulturális szektor ennek megfelelően közösségi, közbenső (non-profit vagy harmadik szektor) és magán üzleti szektorra osztható. Ez a szerkezet nemcsak Nyugat-Európára jellemző, hanem növekvő mértékben a kelet-európai országokra is.

A kultúra döntéshozói nem érzik teljes egészében felelősnek magukat az egész kulturális szektorért, hanem csak annak közösségi alapítású részéért. Növekvő probléma a közösségi kulturális szektorra jutó pénzalapok folyamatos csökkenése. A művészetek és a kultúra intézményesített pénzügyi alapjainak elavult szerkezete egyre több gondot vet fel a döntéshozók számára. A művészetek ún. működési támogatásban részesülnek, de a támogatás a független, az alkotó és az avantgárd művészet számára egyre kevesebb.

Számos kulturális döntéshozó úgy véli, hogy a kulturális ágazathoz tartozó kereskedelmi szórakoztatás üzlet, amely piaci szabályok szerint működik, így nem esik az ő felelősségi körükbe. Még senki nem foglalkozott azzal a kérdéssel, hogy a támogatott művészet vajon sokkal újítóbba jellegű-e, mint a kulturális ágazat más területei. A legtöbb esetben a művészetek termékei, az irodalom és a könyvpiac, a zenei szektor nagy része a magántulajdonú üzleti szférához tartozik.

A művészek megélhetését a közösségi és a magán kulturális szektorok biztosítják. Egy korábbi, a németországi viszonyokat elemző tanulmány arra mutatott rá, hogy a szűkebb értelemben vett művészek (zenészek, képzőművészek, írók, előadóművészek) mintegy kétharmadát a közösségi szektor foglalkoztatta. Mindezek ellenére a kulturális szektor foglalkoztatottjainak többsége – beleértve a nem művészi foglalkozásokat is – a magánszektorban dolgozik.

A legfrissebb foglalkoztatási adatok strukturális változás jeleit mutatják, mely szerint jelentősen nő a művészeti és más foglalkozásuk száma a magánszférában. A közösségi szektor e téren egyre inkább elveszti stratégiai jelentőségét. E változás számos művész megélhetését fenyegeti, mivel teljes munkaidős állásuk a legtöbb esetben részmunkaidőssé változik a magánszférában, és a művészek helyzete így egyre bizonytalanabbá válik.

A téma fontosságát jelzi, hogy a jelenség tanulmányozására egy kísérleti projektet indítottak el. A csoportban bécsi, bonni és zürichi kutatók vesznek



részt. A projektet a Zürichi Művészeti Egyetem vezeti, és célja a kulturális ágazat harmonizált adatbázisának kifejlesztése a közös európai adatbázis számára.

A kultúrának és a médiának döntő szerepe van a különböző európai országok nemzeti, identitásának és regionális változatosságának megőrzésében és fejlesztésében, amit több mint tíz éve hangsúlyoznak. Az 1990-es évek közepétől kezdve az Európai Unió is megvitatta ezt a kérdést és ezzel kapcsolatban több fontos dokumentumot állított össze. Az egyik legfontosabb döntésük a kulturális statisztika támogatása volt, amit 1996 első félévében az Európai Tanács elé terjesztettek. Az Európai Bizottság úgy véli, hogy a versenyképesség, a gazdasági növekedés és a foglalkoztatás csak úgy támogatható, ha a fontos gazdasági és foglalkoztatási adatok gyűjtését kifejlesztik és tökéletesítik.

Jelenleg az EU több szakmai testülete foglalkozik a kulturális szektor harmonizált adatgyűjtésével. Az EUROSTAT kulturális statisztikával foglalkozó munkacsoportja már eddig is értékes előtanulmányokat készített. A Foglalkoztatási Főigazgatóság számos hasznos módszertani és gyakorlati útmutatást adott a munkához, melyek kiindulópontként szolgálhatnak további összehasonlító tanulmányok kidolgozásához, hogy megalkossák a szűkebb értelemben vett kulturális szektor definícióját.

A Gazdasági Tevékenységek Általános Osztályozása (Nomenclature générale des activités économiques dans les Communautés Européennes – NACE) nem tartalmazza a kulturális ágazatokat. Minden egyes eddigi tanulmány a gazdasági szektorok (termelés, kereskedelem, szolgáltatás) saját egyéni újraosztályozását fogalmazta meg. A kutatási terv szorgalmazza az Európai Unió NACE osztályozásának maximális használatát a nemzetközi összehasonlításoknál. A kutatócsoport reméli, hogy sikerül a kulturális ágazatoknak egy olyan osztályozását megalkotni, amely Európa-szerte alkalmazható és átadható az Európai Unión kívüli országoknak is.

A kulturális ágazatok (beleértve a közösségi alapítású szektorokat) mintegy 2,6 millió munkahelyet adnak az Európai Unió 15 országában. Az EU munkaerő felvétele szerint ők teljes és részmunkaidőben foglalkoztatott dolgozók, akiknek fő jövedelmi forrása a kulturális szektor. Ugyanezen felvétellel 1999-ben 6 millió foglalkoztatott mutat ki a mezőgazdaságban, közöttük több mint egy milliót Olaszországban, Németországban és Spanyolországban, míg számuk valamivel egy millió alatt maradt Franciaországban. Politikai jelentőségük ellenére a kulturális szektor foglalkoztatási adatait reálisan értékelni meglehetősen nehéz. A 2,6 millió foglalkoz-

tatott egy szűken értelmezett, illetve nagyon konzervatív módon meghatározott definíción alapul. A Foglalkoztatási Főigazgatóság említett tanulmánya a kulturális ágazatban dolgozók számát 7,2 millió főben alapítja meg.

A legújabb adatok szerint a kulturális foglalkoztatottak arányában meglepő a hasonlóság az Európai Unió tagországaiban. Az Egyesült Királyságban és Németországban egyaránt több mint 600 ezer fő talál munkát a kulturális szektorban. Franciaországban 350 ezer főt, Spanyolországban és Olaszországban megközelítően 200 ezer főt, Hollandiában 166 ezer főt foglalkoztatnak. A többi EU-tagországban 30 és 60 ezer fő közötti az e szektorban dolgozók száma, kivéve a legkisebb országot, Luxemburgot, ahol mindössze 3 ezer fő.

Meglepő, hogy a kulturális szektor foglalkoztatottjainak nemzetközi összehasonlítása bizonyos regionális különbségeket mutat. A kultúrában foglalkoztatottak EU-beli aránya az összes foglalkoztatotthoz viszonyítva 1,7 százalék, mely arányok országonként 0,9 és 2,5 százalék között szóródnak. Minden európai országban, a teljes munkaidőben foglalkoztatottaknak egy viszonylag széles és homogén rétege dolgozik a kulturális szektorban. Ez nem szükségképpen igaz más szolgáltató szektorokra. A sport szektor foglalkoztatási arányaiban nagy különbségek alakultak ki, van ahol hatszoros különbség adódik az országok közötti foglalkoztatási arányokban. A kulturális szektorban ugyanez csak 2,5-szeres, már ami a legmagasabb értékű Dánia (2,5 %) és a legalacsonyabb mutatójú Olaszország (0,9 %) között mérhető.

A növekedési ráták a kulturális szektorban nagyobb különbségeket mutatnak. Az e szektorban foglalkoztatottak aránya az EU-ban 1995 és 1999 között átlagosan 3,8 százalékponttal nőtt, miközben az összes foglalkoztatottaké ugyanezen időszak alatt csak 1,2 százalék volt. Olaszország és Finnország növekedési üteme a legmagasabb, mintegy 8-9 százalék, de átlag feletti volt Írországban, Portugáliában, Hollandiában és Németországban (6-7 %). Luxemburg, Spanyolország, Dánia és Belgium átlagos növekedést ért el (3-5 %), de meglepő módon azokban az országokban, melyeknek kulturális öröksége igen híres (Görögország, Franciaország és Ausztria), a növekedés mindössze 1-2 százalék.

A kutatásban részt vevők projektjének célja egy közös európai adatbázis kifejlesztése. Mindezek elérésére kezdeményezni fogják a kulturális ágazatok összehasonlító tanulmányainak elkészítését más európai országok számára is.

Jelenleg a következő hipotézisek fogalmazhatók meg. 1. A kulturális ágazatok főleg a kis- és közé-

pes- vállalkozásokban foglalkoztatják a dolgozókat. 2. A kulturális ágazatok ellensúlyt képeznek a hagyományos ágazatokkal szemben, amelyek csökkenést mutatnak a strukturális változások hatására. 3. Az e szektorba tartozó vállalkozásokban a munka intenzív jellegű, a munkaerő az átlagosnál képzetesebb, innovatívabb és magasabb szaktudású. 4. Az itt dolgozók között a nők aránya sokkal magasabb, mint a hagyományos ágazatokban. 5. A kulturális ágazatok működése megfelelő környezetet és infrastruktúrát igényel. A hasonló tevékenységet végzők (termelők és szolgáltatók) között az együttműködésre regionális gazdasági hálózat jött létre. 6. A szektorban a helyi jelleg meghatározó, mely összeegyeztethető más életfeltételekkel. A munka és a szabadidő helyi, társadalmi és funkcionális kapcsolatainak új formái fejlődnek ki. 7. Ezen ágazatok akkor lesznek sikeresek, ha ők hozzák létre azokat a kulturális javakat és szolgáltatásokat, melyek eredetiek, ötletgazdagok és nehezen hamisíthatók.

A kutatócsoport célja, hogy hozzájáruljon ahhoz a vitához, ami a kulturális és a gazdaságpolitika integrált megközelítését adja. A kulturális minisztériumok egyre növekvő felelősséget éreznek a média és a kommunikációs szektor iránt. (Az Egyesült Király-

ságban, Franciaországban, Németországban vagy Ausztriában az említett minisztérium neve már: a kultúra és a média és/vagy a kommunikáció minisztériuma.)

E koncepciók alapján nemcsak nemzeti, hanem regionális és helyi szinten is sikeresen fejlesztették ki a kultúra és a gazdaság intézményrendszerét. Az Észak-Rajna-Vesztfáliai tartomány Gazdasági Minisztériuma például teljesen új stratégiák alapján, sikeres együttműködést fejlesztett ki az ottani kulturális minisztériummal. E minisztérium most kísérletezi ki, hogy miként kell támogatni a kezdő művészeket, stúdiókat, a kultúra kis sejtjeit, hogy ne kerüljenek összeütközésbe korábbi támogatási rendszerük szabályaival. Előrelátásra van szükség, hogy kifejlesszék a kultúra izléses és gazdasági szempontokat is figyelembe vevő körét, mely életünk szerves tartozéka. Ehhez támogatni kell a művészeket, a kulturális szektort és képessé kell tenni a régiókat a fejlődés fenntartására. Ennek megvalósítását már 1999-ben megfogalmazták az EU Miniszterek Tanácsa, az Európai Parlament és az Európai Tanács programdokumentumaiban.

(Ism.: *Hajnal Béla*)

## TÁRSADALOMSTATISZTIKA – DEMOGRÁFIA

ROTHSTEIN, D. S.:

### A FIATALKORÚAK FOGLALKOZTATOTTSÁGA AZ EGYESÜLT ÁLLAMOKBAN

(Youth employment in the United States.) – *Monthly Labour Review*, 2001. 8. sz. 6–16. p.

A cikk szerzője, *D. S. Rothstein*, az Egyesült Államok Munkaügyi Hivatala Foglalkoztatás- és Munkanélküliség-statisztikai Irodájának munkatársa azt tűzte ki célul, hogy részletesen ismertesse az egyesült államokbeli fiatalok foglalkoztatottságát egy új típusú „A fiatalok 1997. évi nemzeti longitudinális felmérése” (National Longitudinal Survey of Youth 1997 – NLSY1997) c. felvétel adatainak felhasználásával.

A szerző ismerteti a fiatalok foglalkoztatásának gyakorisági, intenzitás- és időmegoszlási (a tanítási időszak és a nyári szünet közötti) adatait, bemutatja azokat az iparágakat és foglalkozási ágakat, amelyekben a fiatalok általában munkát vállalnak, valamint megvizsgálja a foglalkoztatottságban mutatkozó, nemek, etnikai csoportok, a háztartások bevételére és a családi szerkezet szerinti különbségeket.

A cikkben bemutatott adatok az NLSY1997 első interjúi során végrehajtott, országos szintű reprezentatív adatfelvételtől származnak, mintegy 9000, 1980. január 1. és 1984. december 31. között született fiatalok megkérdezésének eredményeként. Az első interjúra 1997-ben került sor, amikor ezeknek a fiataloknak az életkora 12 és 17 év között volt. A minta interjúalanyait évenkénti gyakorisággal kikérdezik munkaerő-piaci tapasztalataikról, valamint számos egyéb témakörrel kapcsolatban, mint például az iskolai és családi háttérrel összefüggő kérdések. Az ismertetett adatfelvétel célja longitudinális felmérés a fiatalok alkalmazásának tapasztalatairól, nem pedig munkaerő-piaci státusuk megállapítása egy adott időpontban. Ennek érdekében a felvétel adatszolgáltatóitól azt kéri, hogy sorolják fel az összes munkájukat, amelyet 14 éves koruktól kezdve a megkérdezésükig alkalmazottként végeztek. A kérdezőbiztos naplót vezet, amelyet megmutat az adatszolgáltatónak, hogy ellenőrizzék az egyes munkavállalások kezdési és befejezési időpontját, valamint az ezen dátumok közötti olyan időtartamokat, amelyek alatt a válaszadó nem állt alkalmazásban. Ezen kívül a 14 éven felüli adatszolgáltatókat felké-

rik, hogy sorolják fel 14 éves koruktól kezdve az interjú dátumáig folytatott összes szabadúszó munkavégzésüket is. Itt is napló segítségével igazolják a munkakezdés és -befejezés időpontját. A szabadúszóként végezhető munkák szórványos jellege miatt azonban az egyes dátumok közötti munkaszüneteltetések időtartamának adatait nem gyűjtik. A válasszadók megadják az összes szabadúszóként végzett munkájuk jellemző adatait is.

A tanulmányban közölt táblák a fiatalok foglalkoztatási adatait tartalmazzák mind alkalmazottként végzett munka, mind szabadúszóként, munka szerinti bontásban, egy meghatározott időtartamon belül, 14 és 15 éves korban. A felmérés a fiatal 14. és 15. születésnapja közötti teljes 52 hetes időtartam valamely időszakában végzett munkát tekinti 14 éves korban történt foglalkoztatásnak. A fiatalok foglalkoztatásának idejét tartalmazó táblák és grafikonok az 1996-os naptári évre és 1981-re mint születési évre vonatkoznak.

Más felvételektől eltérően a felmérés foglalkozik az ún. „nagyon fiatalok” foglalkoztatottsági adataival is: a 12 vagy 13 évesek 12 éves koruktól kezdve végzett munkájáról közöl adatokat. Erre az életkorra vonatkozóan nem tesz különbséget az alkalmazottként vagy a szabadúszóként végzett munka között.

A korábbi kutatások szerint a fiatalok munkavállalási magatartását több tényező befolyásolja aszerint, hogy melyik nemhez tartoznak, milyen etnikumból származnak, mennyi a háztartásuk bevétele, és milyen családi szerkezetben élnek. A cikkben közölt táblák is e tényezők szerinti bontásban közlik a fiatalok foglalkoztatottsági adatait.

A továbbiakban a fiatalok foglalkoztatottságának gyakoriságát vizsgálva megállapítható, hogy a felmérés adatai szerint a 14-15 éves korú fiatalok jelentős hányada áll alkalmazásban. Több mint 57 százalékuknak már 14 évesen volt munkahelye. Többeségük ebben az időszakban szabadúszóként dolgozott. A 15 évesek 64 százaléka állt alkalmazásban legalább egy bizonyos időszakban, de ők már inkább alkalmazottként, és nem szabadúszóként dolgoztak. A 14 éves korosztályban a fiúk 55, a lányok 59 százaléka vállalt munkát, míg a 15 éveseknél szinte egyenlő ez az arány (63%, 64%). A szabadúszóként végzett munka és az alkalmazásban végzett munka közötti megoszlás jelentősen eltér a nemektől függően mindkét korosztályban: a lányok inkább szabadúszó, a fiúk alkalmazotti munkaterületen tevékenykedtek nagyobb létszámban.

A fejezetből az is kiténik, hogy ez a felmérés azt a korábbi kutatási eredményt is igazolta, hogy lényeges eltérés mutatkozik a fiatalok foglalkozta-

tottságának gyakoriságában a bőrszín és az etnikai hovatartozás szerint. A fehérek aránya jóval magasabb az afro-amerikai vagy spanyol eredetűek arányánál ebben a korosztályban. A 14 évesek között a fehérek 64, az afro-amerikaiak 43 és a spanyolok 41 százaléka dolgozott, míg a 15 éves korosztályban a fehérek 72 százaléka, a spanyolok 48 százaléka. Az afro-amerikaiak aránya azonban alig változott (44 %).

A háztartások jövedelme szerinti megoszlás figyelemre méltó eredményt mutat: az alacsonyabb jövedelmű háztartásokban (évi 25 ezer dollár alatt) a 14 éves fiatalok elhelyezkedési aránya jóval alacsonyabb (49 %), mint a magasabb jövedelmű háztartásokban (63 %). A 15 éves korosztályban az elhelyezkedés gyakorisága között hasonló eltérés mutatkozik (52 % és 69 %).

A családszerkezet szerinti vizsgálatból kiténik, hogy a mindkét szülővel rendelkező családok gyermekei 14 éves korukban sokkal nagyobb százalékban vállaltak munkát, mint ahol a szülői szerepet csupán az anya töltötte be. A 15 éves korosztály adatai ebben a kategóriában nem mutatnak jelentős különbséget.

A munkaheteket véve alapul, a fiatalok alkalmazottként végzett munkavállalásának gyakorisága a 14 éves korosztályban 24 százalékos, a 15 éves korosztályban 38 százalékos értéket mutatott, ami lényeges emelkedést jelent. Ebben a kategóriában a fiúk mindkét korosztályban jóval nagyobb arányban vállaltak munkát, mint a lányok. Ugyanígy a fehérek elhelyezkedési aránya a két korosztályban jelentősen meghaladta mind az afro-amerikai, mind a spanyol eredetű fiatalokét.

Az iparágak és a foglalkozási ágak szerinti munkavállalási adatok azt mutatják, hogy a 14-15 éves, alkalmazásban álló amerikai fiatalok csaknem 60 százaléka mindössze tíz foglalkozási ágban helyezkedett el, ezek közül különösen kettőnek jut kiemelt szerep: az étkezdékben és az italt felszolgáló helyeken kínált álláslehetőségek. A fiúk jelentős hányada vállalt az építőiparban munkát, míg a fiatal lányok körében jelentős a magánháztartásoknál elhelyezkedők száma.

A szabadúszóként végzett munkák sorában első helyen a lányoknál a kisgyermekgondozás (babsitterek), fiúknál a kerti munkák állnak.

Az iskoláskorúak munkavállalási adataiból kiténik, hogy 1966-ban a legtöbb fiatal a nyár közepén vállalt munkát, a legkevésbé tavaszi szemeszter idején, ennél valamivel többen az őszi szemeszterben. Ebben az évben a fiúk elhelyezkedési aránya magasabb volt, mint a lányoké.

A legfiatalabbak foglalkoztatottságát bemutató tábla szerint a 12 éves korosztályba tartozóknak

csaknem a fele dolgozott a vizsgált időszakban, ezek több mint 50 százaléka babysitterként, 40 százaléka pedig kerti munkát vállalt.

A tanulmány szerzője megállapítja, hogy a felmérés szerint a 14-15 éves korosztályban a fiatalok inkább alkalmazottként, mint szabadúszóként vállaltak munkát a vizsgált időszakban. Mindkét korosztályban jelentős eltérések mutatkoznak a nemek és a børszín szerinti megoszlásban, és mindegyikben in-

kább a lányok választják a szabadúszó munkavállalási formát. Végül, a fehérek sokkal nagyobb arányban helyezkednek el, mint az afro-amerikaiak vagy a spanyol eredetűek. A fiatalok munkavállalásának időmegoszlását tekintve a vizsgálat kimutatta, hogy az alkalmazottként dolgozó 14-15 éves korosztály a tanítási időszakban is vállal munkát.

(Ism.: *Hejna Ferencné*)

## KÜLFÖLDI FOLYÓIRATSZEMLE

### Allgemeines Statistisches Archiv

A NÉMET STATISZTIKAI TÁRSASÁG  
FOLYÓIRATA

2001. ÉVI 4. SZÁM

*Rietzler, K. – Stephan, S. – Wolters, J.:* Aggregálás és szezonális kiigazítás: empirikus eredmények az EMU negyedéves nemzeti számlái esetén.

*Jung, R.C. – Liesenfeld, R.:* Esetszámok idősor modellbecslése hatékony fontossági mintavétellel.

*Biewen, M.:* Részleges nemváltás és egyenlőtlenségmérés: bizonyíték a német kereseti eloszlásból.

*Eichwede, R. – Krumbholz, W.:* ASN-optimális kettős mintavételi tervek minőségi változókra.

*Abberger, K.:* Lokális korreláció közvetítő változóval.

*Kladroba, A.:* Mi az új az adatbányászásban? Egy megjegyzés az „új” adatelemzéshez statisztikai nézőpontból.

*Krud, W. – Ernst, N.:* A magánháztartások gazdasági számításainak újragondolásáról.



A FRANCIA GAZDASÁGI  
ÉS PÉNZÜGYMINISZTERIUM ÉS A STATISZTIKAI  
ÉS GAZDASÁGKUTATÓ INTÉZET FOLYÓIRATA

2001. ÉVI 3. SZÁM

*Bascher, J. – Niedergang, M.:* A közkiadások egy újabb megbontása.

*Dormont, B. – Fourgère, D. – Prieto, A.:* Az egyszerű depresszív segély hatása a munkába való visszatérésre.

*Louvot-Runavot, C.:* Lakásépítés az Európai Unióban: a tulajdonlás utoléri a bérlést.

*Barnet-Verzat, C. – Wolff, F.-C.:* Gyermek zsebpénze: az önmálló pénzközelés elsajátítása.

2001. ÉVI 4. SZÁM

*Gumbert, S. – Vallat, J. C.:* Tőzsdei opciós adózás: nemzetközi áttekintés.

*Baccaini, B.:* Belső vándorlás Franciaországban 1990-től 1999-ig: a nyugat hívása.

*Cases, C. – Misségué, N.:* Markáns állássegimentáció a szolgáltatási szektorban.

*Seguin, S.:* A nagy specializált áruházak átveszik a hipermarketeket.

2001. ÉVI 4. SZÁM

*Rivière, C.A.:* A telefon: a társadalmi integráció eleme.  
*Rouquette, C.:* Szabadságra menni: állandósulnak az egyenlőtlenségek.

*Gilbert, L. – Kamionka, T. – Lacroix, G.:* A kormány foglalkoztatási programjának hatásai fiatal hátrányos helyzetű férfiaknál Quebecben.



A BIRMINGHAMI EGYETEM FOLYÓIRATA

2001. ÉVI 3. SZÁM

*Bulatov, A. S.:* Orosz működőtőke-befektetés külföldön: történelem, motivációk, pénzügyi ellenőrzés és tervezés.

*Sun, L.:* Árolló, sorbanállás és korlátozás: kiterjesztett keret az egyszerű szocialista felhalmozás megértéséhez.

*Estrin, S. – Lazarova, S. – Urga, G.:* Konvergencia az átmeneti országokban – összpontosítás a beruházásra: Közép- és Kelet-Európa, 1970–1996.

*Revolteila, D.:* A vállalatok finanszírozása Csehországban: adósság- és cégspecifikus változók.



A NEMZETKÖZI STATISZTIKAI INTÉZET  
FOLYÓIRATA

2001. ÉVI 3. SZÁM

*Agresti, A. – Natarajan, R.:* Klaszterezett rendezett kategóriás adatok modellezése: áttekintés.

*Molina, E. A. – Smith, T. M. F. – Sugden, R. A.:* A túlszóródás modellezése komplex felvételi adatok esetén.

*Hand, D. J. – Keming, Y.:* Az idióta Bayes – végül nem is olyan buta ?

*Lecoutre, B. – Lecoutre, M. P. – Poitevineau, J.:* A szignifikancia-tesztek félrevezető és helytelen használata: elkerülhetetlen lesz a bayesi választás ?

*Holst, E. – Thyregod, P. – Wilrich, P. Th.:* A konformitási tesztől és a kétlépcsős eljárások használatáról.

*Gupta, R. D. – Richards, D.:* A Dirichlet- és Liouville-eloszlások története.

*Ayhan, H. Ö.:* Nemek szerinti statisztikák: intézkedések a nemeken belüli torzítás csökkentésére a mezőgazdasági felvételekben.

*Singh, M. P. – Hidiroglou, M. A. – Gambino, J. G. – Kovačević, M. S.:* Becslési módszerek és kapcsolódó rendszerek a Kanadai Statisztikai Hivatalnál.

*Hidiroglou, M.A. – Laniel, N.:* Mintavételi és becslési kérdések éves és évközi kanadai vállalati felvételeknél.



A SVÉD KÖZPONTI STATISZTIKAI HIVATAL  
FOLYÓIRATA

2001. ÉVI 3. SZÁM

*Pickery, J. – Loosveldt, G.:* Azoknak a kérdésjellegzetességeknek a feltárása, amelyek a kérdezőbiztosító hatásokat közvetítik a részleges nemválaszolók csökkentésére.

*Blasius, J. – Thiessen, V.:* Közömbös válaszok használata felvételi kérdésekben: a többszörös korrespondencia-analízis alkalmazása.

*Bradley, R.:* Véges mintahatások a helyettesítési torzítás becslésében a fogyasztói árindexnél.

*Longford, N. T.:* A foglalkoztatottság arányainak és összetételének becslése a norvég önkormányzatoknál.

*Moriarity, C. – Scheuren, F.:* Statisztikai párosítás: egy paradigma az eljárási bizonytalanság értékeléséhez.

*Kadane, J. B.:* Néhány statisztikai probléma adatállományok összefűzésénél.



AZ ANGOL KIRÁLYI STATISZTIKAI TÁRSASÁG  
FOLYÓIRATA (A SOROZAT)

2001. ÉVI 3. SZÁM

*Simpson, L.:* Az Egyesült Királyság 2001-es cenzusa – előrelépés és ígéret.

*Dale, A. – Elliot, M.:* Javaslatok anonim rekordok 2001-es mintáira: a felfedés kockázat értékelése.

*Lipsitz, S. R. – Williamson, J. – Klar, N. – Ibrahim, J. – Parzen, M.:* Egy egyszerű módszer regressziós modell esetén egy aránypár közötti becslésére.

*Lloyd, C. J. – Yip, P. S. F.:* Öngyilkossági módok összehasonlítása Ausztrália és Hong Kong esetében.

*Leung, D. H. Y.:* Statisztikai módszerek klinikai vizsgálatokhoz helyettesítő végpontok jelenlétében.

*Thompson, M. L. – Edland, S. D. – Gibbons, L. E. – McCurry, S. M.:* Referencia-tartományok becslése rétegzett kétlépcsős mintákból.

*Forster, M. – Jones, A. M.:* A dohányzók szerepe a dohányzás elkezdésében és abbahagyásában: brit adatok időtartam-elemzése.

*Ali, M. M. – Marshall, T. – Babiker, A. G.:* Nem teljes időtartamok elemzése a fogamzásgátló szerek használatára alkalmazva.

## POPULATION

A FRANCIA DEMOGRÁFIAI INTÉZET FOLYÓIRATA

2001. ÉVI 4. SZÁM

*Hammer, R. – Forster, M. – Jones, A. M.:* Rokoni kapcsolatok a fiatal svájci generációkban: családi ág, struktúra és funkció.

*Charton, L. – Wanner, P.:* Az első párkapcsolat kialakulása Svájcban: az együttélés típusának kiválasztása és trendek a házasságon kívüli együttélésben.

*Prioux, F.:* Franciaország jelenlegi demográfiai helyzetének alakulása.

*Toulemon, L. – Mazuy, M.:* A gyermekvállalás késleltetett, de a termékenység stabil.

*Toulemon, L. – Mazuy, M.:* Stabil viselkedési hipotézisen alapuló öt termékenységi előrejelzés.

*Razamifanjato, J. és mások:* Madagaszkár demográfiai helyzete.



A SZLOVÁK STATISZTIKAI HIVATAL  
FOLYÓIRATA

2001. ÉVI 2. SZÁM

*Baláz, P.:* AZ ESNA megvalósítása Szlovákia éves nemzeti számláiban.

*Luhová, M.:* Rövid áttekintés a szektorális elszámolásokról.

*Jakúbeková, J.:* A negyedéves nemzeti számlák kialakításának helyzete és kilátásai Szlovákiában nemzetközi összefüggésben.

*Karol'ová, V.:* Beszámolási adatok a közigazgatás deficitjéről és adósságáról.

*Kolesárová, L.:* A koordinálással és információszolgáltatással kapcsolatos tapasztalatok a nemzeti számla rendszerben.

*Kosseyová, O.:* Harmonizált fogyasztói árindexek.

*Varholová, R.:* A termelőiár-statisztika céljai 2001 és 2003 között.

*Cár, M.:* Szlovákia és Európa válogatott társadalmi-gazdasági jellemzők fényében.

2001. ÉVI 3. SZÁM

*Marcek, D.:* Rövid távú idősorok modellezése szezonális komponenssel, strukturális modellek felhasználásával.

*Mészáros, J.:* A halálózásíráta-összehasonlítások lehetőségei különböző korstruktúrákban.

*Chajdiak, J.:* Az élveszületések számának becslése Szlovákiában 2020-ig.



AZ EGYESÜLT ÁLLAMOK  
MATEMATIKAI STATISZTIKAI INTÉZETÉNEK  
FOLYÓIRATA

2001. ÉVI 2. SZÁM

*Brown, L. D. – Cai, T. T. – DasGupta, A.:* Intervallum-becsülés egy binomiális arányra.

*Opsomer, J. – Wang, Y. – Yang, Y.:* Nemparaméteres regresszió korrelált hibákkal.

*Abt, M. és mások.:* Egy szekvenciális módszer ölomösszetevők azonosítására nagy kémiai adatbázisokban.

*Hinkelmann, K.:* Emlékezés Oscar Kempthorne-ra (1919–2000).

*Olshen, R.:* Beszélgetés Leo Breimannel.



AZ EGYESÜLT NEMZETEK EURÓPAI GAZDASÁGI  
BIZOTTSÁGÁNAK FOLYÓIRATA

2001. ÉVI 1. SZÁM

*Oudhof, K.:* A GDI mint a fejlődés nemek szerinti jellemzőinek mérési eszköze az EGB-régióban.

*Keuzenkamp, S.:* Nők a politikai és társadalmi döntéshozatalban Hollandiában.

*Kjeldstad, R. – Kristiansen, J. E.:* Regionális nemiegyenlőségi-index készítése: reflexiók a norvég adatokon alapuló első tapasztalatokra.

*Hoffmann, E. – Greenwood, A. M.:* Statisztikák munkaidő megállapodásokról.

*Ward, M.:* Nemzetközi árszintek és globális infláció.

*Heston, A. – Summers, R. – Aten, B.:* Árstruktúrák, minőségi tényező és láncolás.

*Mills, M. – Trovato, F.:* A terhesség hatása a házasságkötésre az élettársi kapcsolatban Kanadában, Hollandiában és Lettországban.

*Karlsson, J.:* Minőség szerint javított árindex az ipari robotokra.

2001. ÉVI 2–3. SZÁM

*Malaguerra, C.:* Bevezetés.

*Hammarberg, T.:* Az igazság keresése: szükség van az emberi jogok figyelésére mérhető és megbízható eszközökkel.

*Omar, E.:* Fejlődés és emberi jogok: a nemzetközi közösség növekvő statisztikai igénye.

*Mokhiber, C. G.:* A méltóság mérése felé: mutatók a jogalapú fejlődéshez.

*Ball, P.:* Egy álláspont kifejtése: a statisztika szerepe az emberi jogok beszámolójában.

*Kucera, D.:* Az alapvető munkajogok mérése.

*Ashagrie, K.:* Új módszerek és megújult felvételek a gyermekek kizsákmányolásának méréséhez.

*Goldman, G.:* A kisebbségek meghatározása és megfigyelése: objektív értékelés.

*Raines, M. D.:* Az együttműködés megnyerése a többkultúrú válaszadók közösségében: az amerikai Census Bureauak az újonnan bevándorolt népesség számbavételét célzó erőfeszítéseinek áttekintése.

*Brunborg, H.:* A statisztikai elemzés hozzájárulása a nemzetközi büntető bíróságok vizsgálataihoz.

*Fukuda-Parr, S.:* Az emberi fejlődés és szembeni jogok mutatói – átfedések, különbségek... és mi van az emberi fejlődés indexével?

*Guzman, M. M.:* Az események vizsgálata és dokumentálása mint módszertan az emberi jogok megsértésének figyelésében.

*Thede, N.:* Emberi jogok és statisztika: néhány reflexió a „nem ember földjén” fogalom és mutató között.

*Mottet, C. – Suarez de Miguel, R.:* A fejlődés és emberi jogok figyelése? Egy projekt a kihívás kezelésére.

## Statistische Nachrichten

AZ OSZTRÁK KÖZPONTI STATISZTIKAI HIVATAL  
FOLYÓIRATA

2001. ÉVI 9. SZÁM

Népességelőrejelzés Ausztriára és a tartományokra, 2001–2050.

Magánháztartások felszereltsége: 1999/2000-es háztartási költségvetés felvétele.

Ausztria társadalombiztosítási intézményei 2000-ben.

A feldolgozó iparágak 1999-es teljesítmény és szerkezeti felvétele.

Termelői árindex (1996 = 100): hibrid index számítása.

Az osztrákok szabadidős utazásai 2000-ben: 2000. decemberi mikrocenzus.

Integrált bér és jövedelemadó 1998-ban.

2001. ÉVI 10. SZÁM

A 2001-es népszámlálás: előzetes eredmények településként.

A 2000-es demográfiai struktúra és trendek Ausztriában.

A beiratkozott tanulók számának alakulása, 1990/91–2000/01.

A magánháztartások kiadásai – regionális eredmények; a 1999/2000-es háztartási költségvetés felvétele.

Építési költségek és finanszírozás 1999-ben.

Lakásépítés 2000-ben.

Állati termékek szállítási mérlege, 2000.

Öko-adók, 1997–2000.

A háztartások energiafogyasztása 1999/2000-ben: a 2000. júniusi mikrocenzus.

Építési költség indexek (2000=100); a mélyépítési ágazatok felülvizsgálata.

Külkereskedelem 2001. januártól júniusig; előzetes eredmények.

2001. ÉVI 11. SZÁM

A 2001-es népszámlálás: előzetes eredmények településméret és városi régiók szerint.

Foglalkoztatottság és munkanélküliség járasként 2001. július végén.

Nettó háztartási bevételek, 1999/2000.  
Teljes fenyőkitermelés 2000-ben.  
Vadászati statisztika, 2000/2001.  
Környezetvédelmi kiadások Ausztriában 1999-ben.  
Áruszállítás a Dunán 2000-ben.  
Állami pénzügy.  
Nemzeti számlák az ESA 95 szerint: a 2000. év eredményei.  
Forgalmiadó-statisztika, 1998.

## 2001. ÉVI 12. SZÁM

A 2001-es népszámlálás: előzetes eredmények életkor és nemek szerint.

Háztartás és család előrejelzés tartományonként 1991-től 2030-ig; 2001 újraszámítása.

A magánháztartások felszereltsége – társadalomstatisztikai eredmények; 1999/2000-es háztartási költségvetés felvétel.

Lakások fűtése 2001-ben; 2001. júniusi mikrocenzus. Termények és szántóföld 2001-ben.

A tartományi kormányzatok energiafogyasztása 2000-ben.

A 2000. évi anyagfelhasználás-felvétel a feldolgozó iparágakban.

Osztrák idegenforgalmi szatellit elszámolások – közgazdaságtan, módszertan és eredmények.

Osztrák vállalatok közötti áruszállítása 2000-ben.

Végso fogyasztás a nemzeti számlákban.

Egészségügyi kiadások 1995-től 2000-ig.

Nemzeti számlák az ESA 95 szerint: 2000 fő eredményei.

**statistika**  
EKONOMICKO - STATISTICKÝ ČASOPIS

A CSEH STATISZTIKAI HIVATAL  
FOLYÓIRATA

## 2001. ÉVI 10. SZÁM

*Pilková, J. – Jileček, A.:* A cseh háztartások jövedelmeinek árának és fogyasztásának trendjei, a 90-es években.

*Ondrus, V.:* A nonprofit intézményi szektor a nemzeti számlák oldaláról nézve.

*Petrikovits, E.:* Közigazgatási regiszterek – koncepció és törvényi háttér.

*Siroky, J.:* Regionális statisztikák a Cseh Statisztikai Hivatal adatbázisaiban – regionális és körzeti adatok.

*Fischer, J.:* Alternatív módszerek a munkanélküliség mérésére.

*Hochmaulová, D.:* Betegbiztosítási szolgáltatások változása 1995 óta.

Csehország gazdasága: fő mutatók.

## 2001. ÉVI 11. SZÁM

*Kunc, Z. – Vlášek, J.:* Az építőipari vállalatok pénzügyi teljesítménye 1993 és 2000 között.

*Zeleny, M.:* A jövedelmi szegénység Csehországban, EU-módszertan felhasználásával.

*Kociánová, S.:* Az EU-előírások teljesítése a termelői árak területén.

*Drápal, S.:* Régiók nemzetközi statisztikai összehasonlítása.



AZ OROSZ ÁLLAMI STATISZTIKAI  
BIZOTTSÁG FOLYÓIRATA

## 2001. ÉVI 9. SZÁM

*Gulidov, A. G. – Golovanov, Yu. K. – Sljusarenko, A. T.:* Információs kilátások és az Orosz Statisztikai Hivatal számítógépes rendszerének fejlesztése.

*Medvedev, B. G. – Chistjakov, E. G.:* A föderáció társadalmi és gazdasági fejlődésének statisztikai elemzése és előrejelzése.

*Kuznecova, V. E. – Sivel'kin, V. A. – Mkhitarjan, V. S.:* Makroökonómiai mutatók kapcsolatának vizsgálata egy régió esetén.

*Khasaev, G. R. – Cybatov, V. A.:* Egy kombinált pénzügyi mérleg kidolgozásának és elemzésének tapasztalata.

*Kuskov, E. A.:* Különbőségek becslési módszertana Oroszország régióinak adózására.

*Nivorozhkina, L. I. – Nivorozhkin, E. M.:* Egy nagyváros munkaerőpiacán regisztrált munkanélküliség időtartamának ökonometriai elemzése.

*Zhandarov, A. M. – Shiller, F. F.:* Statisztikai módszer a társadalmi-gazdasági helyzet változásainak a választói viselkedésre gyakorolt hatásának becslésére.

*Rafikova, N. T.:* Összehasonlítható értékmutatók számitásáról.

*Mikhailova, T. M.:* Statisztikai mutatók időegységeiről.

*Mikhailova, T. M. – Juzbashev, M. M.:* Egy mutató jellemzőinek halmaza.

A társadalmi és gazdasági fejlődés fő mutatói az Orosz Föderációban 1996-2001-ben.

A 2002-es oroszországi népszámlálás.

**Wirtschaft und Statistik**

A NÉMET SZÖVETSÉGI STATISZTIKAI HIVATAL  
FOLYÓIRATA

## 2001. ÉVI 8. SZÁM

A NACE Rev. 1 gazdasági tevékenységgel foglalkozó válogatott ágazatainak homogenitása.

A szálloda- és vendéglátóipar alakulása 2000-ben.

A német marhahús külkereskedelme a BSE- és FMD-krízis hátterében.

Vállalatok a szállítási szektorban – a szállítási tevékenység felvételei és struktúrái.

Közúti utasszállítás 1991-től 2000-ig.

Az életlen át tartó tanulás statisztikai lefedése.

Súlyosan hátrányos helyzetű személyek, 1999.

## 2001. ÉVI 9. SZÁM

Nemzeti számlák, 2001. első félév.

A vásárlóerő-paritás új számítása a fogyasztói árak nemzetközi összehasonlításának hatókörében.

Csődveszteségek, 1996–1998.

Német külkereskedelem Amerikával.

Tengeri szállítás, 2000.  
Egészségre vonatkozó kérdések.

## 2001. ÉVI 10. SZÁM

Minőség és felhasználók.  
Német külkereskedelem, 2001. első félév.  
Gyermekeket is érintő közúti balesetek, 2000.  
Építési célú megtakarítások, 2000.  
Állami pénzügyek, 2001. első félév.  
Az Európai Statisztikai Rendszer (ESS) minősége.

## 2001. ÉVI 11. SZÁM

Környezetgazdasági elszámolások, 2000.  
A népesség-előrejelzések pontossága.  
Nehézségek és a kezelés megoldásai a statisztikai vállalkozói regiszterrel összefüggésben adózási célokkal létrehozott társaságok integrált csoportjainál.  
Az EU-tagállamok kereskedelmi statisztikája.  
Munkanélküliek háztartásai.  
Bruttó állóeszköz-felhalmozás az építőiparban a nemzeti számlákkal összefüggésben.

---

Közzöljük kedves Olvasóinkkal, hogy a *Statisztikai Szemle* májusi és júniusi száma összevontan júniusban jelenik meg.

---