

Statisztikai Szemle

A KÖZPONTI STATISZTIKAI HIVATAL
TUDOMÁNYOS FOLYÓIRATA

SZERKESZTŐBIZOTTSÁG:

DR. BAGÓ ESZTER, DR. BELYÓ PÁL, DR. FAZEKAS KÁROLY, DR. HARCZA ISTVÁN,
DR. JÓZAN PÉTER, DR. KARSAI GÁBOR, DR. LAKATOS MIKLÓS (főszerkesztő), NYITRAI FERENCNÉ DR.,
DR. OBLATH GÁBOR, DR. PUKLI PÉTER (a Szerkesztőbizottság elnöke), DR. RAPPAI GÁBOR,
DR. ROÓZ JÓZSEF, DR. SPÉDER ZSOLT, DR. SZÉP KATALIN, DR. SZILÁGYI GYÖRGY

86. ÉVFOLYAM 4. SZÁM

2008. ÁPRILIS

*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!

ISSN 0039 0690

Megjelenik havonta egyszer
Főszerkesztő: dr. Lakatos Miklós
Osztályvezető: Dobokayné Szabó Orsolya
Kiadja: a Központi Statisztikai Hivatal
A kiadásért felel: dr. Pukli Péter
2008.060 – Xerox Magyarország Kft.

Szakreferensek: Farkas János (társadalomstatistika),
dr. Hajdu Ottó (módszertan), Laczka Sándorné dr. (gazdaságstatistika)
Szerkesztők: Polyák Andrea, Visi Lakatos Mária
Tördelőszerkesztők: Bartha Éva, Simonné Káli Ágnes

Szerkesztőség: Budapest II., Keleti Károly utca 5–7. Postacím: Budapest, 1525. Postafiók 51.
Telefon: 345-6908, 345-6546 Telefax: 345-6594

Internet: www.ksh.hu/statszemle

E-mail: statszemle@ksh.hu

Kiadó: Központi Statisztikai Hivatal, Budapest II., Keleti Károly utca 5–7.

Postacím: Postafiók 51. Budapest, 1525. Telefon: 345-6000

Előfizetésben terjeszti a Magyar Posta Rt. Hírlap Üzletág (1008 Budapest, Orczy tér 1).

Előfizethető közvetlen a postai kézbesítőknél, az ország bármely postáján,
valamint e-mailen (hirlapelofizetes@posta.hu) és faxon (303-3440).

További információ: 06-80-444-444

Előfizetési díj: fél évre 3000 Ft, egy évre 5400 Ft

Beszerezhető a KSH Könyvesboltban. Budapest II., Fényes Elek u. 14-18. Telefon: 345-6789

Tartalom

Tanulmányok

Statisztikai módszerek alkalmazása a fejkvóta számítására a magyar egészségügyben – <i>Nagy Balázs – Rakonczai Pál – Dr. Gulácsi László</i>	321
Esettanulmány egy nyugdíjas házaspár anyagi helyzetének alakulásáról 1990-ről 2007-ig – <i>Dr. Csepely-Knorr András</i>	342
Hiányzó adatok és kezelésük a statisztikai elemzésekben – <i>Oravecz Beatrix</i>	365

Műhely

Mutatók és jobb politikaalkotás: a fenntartható fejlődés kérdésköre – <i>Inna Šteinbuka – Pascal Wolff</i>	385
--	-----

Fórum

Hírek, események	394
------------------------	-----

Szakirodalom

Könyvszemle

Dr. Nemes Erzsébet (szerk.): Fényes Elek breviárium – <i>Marton Ádám</i>	397
--	-----

Folyóiratszemle

Clements, K. W. – Izan, I. H. Y. – Selavanathan, E. A.: Sztochasztikus indexszámok: áttekintés – <i>(Marton Ádám)</i>	401
Evertsson, M.: Nemek szerinti reprodukció – <i>(Tűű Lászlóné)</i>	404
A férfiak és a nők számítógép- és internethasználatát jellemző különbségek – <i>(Lakatos Judit)</i>	406
Kiadók ajánlata	408
Társfolyóiratok	411

Statisztikai módszerek alkalmazása a fejkvóta számítására a magyar egészségügyben*

Nagy Balázs,

a BCE tanársegédje

E-mail: balazs.nagy@uni-corvinus.hu

Rakonczai Pál,

az ELTE tanársegédje

E-mail: paulo@math.elte.hu

Dr. Gulácsi László,

PhD, a BCE habilitált egyetemi docense

E-mail: laszlo.gulacsi@uni-corvinus.hu

A tanulmány célja, hogy egy rövid elméleti áttekintésen és egy gyakorlati példán keresztül bemutassa az egészségügyi forrásallokációban használatos fejkvóta kialakításának sajátosságait Magyarországon. Az első rész a fejkvótakészítés módszertani kérdéseit tekinti át különös tekintettel a fejkvótamodellek építésének folyamatára; a második részben egy betegcsoport elemzésén keresztül bemutatásra kerül a fejkvótakészítés egyik alkalmazási területe. Az elméleti áttekintés és a gyakorlati példa egyaránt rávilágít arra, hogy a finanszírozási formula elkészítéséhez a matematikai-statisztikai eljárások magas szintű ismerete mellett számos, más tudományterület tapasztalatainak együttes felhasználása a célravezető.

TÁRGYSZÓ:

Egészségügyi statisztika.

Matematikai modell.

* A szerzők külön köszönetet mondanak *Gilicze Lászlónak*, az MKB kockázati elemzőjének a regressziós elemzések és az iteratív vizsgálatok módszertanának kidolgozásában nyújtott közreműködéséért.

Az egészségügyi forrásallokációban használatos fejkvóta egy előre meghatározott fix összeg, amelyet az ellátásra jogosult személyek után, meghatározott szolgáltatásokért, meghatározott időszakra fizetnek. A fejkvóta minden érintett után egy adott összeggel számol, amelyet az adott ellátást végző, vagy az ellátások megszervezéséért és finanszírozásáért felelős egészségügyi szervezeteknek fizetnek ki. Az elmúlt húsz évben a fejkvótával történő forráselosztás jelentősége a fejlett országok többségében drámai mértékben megnőtt. Az 1970-es évektől Angliában, az 1980-as évek közepétől az Egyesült Államokban és az 1990-es évektől pedig Nyugat-Európa számos országában (Hollandia, Németország, Belgium, Svédország), illetve Ausztráliában és Új-Zélandon is fontos egészségpolitikai kérdéssé vált (*Rice–Smith* [2001]). Ennek oka, hogy a folyamatosan növekvő egészségügyi költségek megfékezése érdekében az egészségügyi források szétosztására egyre több helyen zárt, előre meghatározott költségvetéseket alkalmaznak. A fejkvóta pedig a zárt, előre meghatározott költségvetések kialakításának átlátható, számon kérhető és jól kommunikálható módszere, amely nemcsak az egészségügyi közgazdászok elvárásainak felel meg, de elfogadható a rendszer többi szereplője számára is (*Milgrom–Roberts* [1990]).

Magyarországon a 2009. január 1-jére tervezetten induló új egészségbiztosítási rendszer egyik alapeleme a fejkvótás forrásallokáció lesz. A T/4221 számú törvényjavaslat szerint az egészségbiztosítási pénztáraknak juttatott megközelítőleg 1100 milliárd forint elosztásáról ennek a módszernek kell gondoskodnia (*Nagy–Sipos–Nagy* [2007], *Magyar Köztársaság Kormánya* [2007]). A tét nem kevés és a feladat nem egyszerű, mivel a nemzetközi tapasztalatok szerint a fejkvóta kialakítása egy igen összetett társadalmi, tudományos, politikai és etikai szempontok figyelembe vételét együttesen igénylő multidiszciplináris folyamat (*Van de Ven–Ellis* [2000]). Feltehetően az egészségügyi rendszer működésének ismeretét, a célok és az eszközök összehangolásának képességét, az egészségpolitikai kontextus pontos megértését, orvosi, közgazdasági, szociológiai, demográfiai és nem utolsósorban matematikai-statisztikai ismereteket (*Nagy* [2006]). A következő áttekintés a matematikai-statisztikai módszertan szerepét vizsgálja először egy rövid elméleti áttekintés majd egy gyakorlati példa bemutatásán keresztül.

1. A fejkvótaformula kialakításának módszerei

A módszerek közül elsőként a kockázatiigazítás módszerét ismertetjük.

1.1. A kockázatkiigazítás

Az egészségügyben hatékonysági és méltányossági megfontolások miatt a finanszírozott szervezeteknek juttatott fejkvótaösszegeket a populáció adottságainak minél pontosabb figyelembe vételével kell meghatározni; köznyelven „súlyozni” szükséges a fejkvótát. Ez a folyamat az ún. *kockázatkiigazítás* (risk adjustment). A probléma matematikailag a következőképpen formalizálható: adott i szervezet B_i költségvetése,

amely a $B_i = \sum_{j=1}^{N_i} c(x_j^i)$ összefüggéssel határozható meg, ahol N_i az emberek létszá-

ma, akiknek az ellátásáért az i szervezet felelős, $c(\cdot)$ a fejkvóta formula és x_j^i a j személy tulajdonságainak vektora; továbbá az egyes B_i szervezetek költségvetésének nagyságát zárt költségvetési keret korlátozza, tehát az összköltségvetés nem lépheti túl az előre meghatározott T keretet, azaz $\sum B_i \approx T$. Mindezeknek megfelelően a kockázatkiigazítás során felmerülő két legfontosabb feladat, hogy a költséget legnagyobb mértékben befolyásoló tulajdonságokat (x) meghatározzuk, illetve az, hogy a meghatározott tulajdonságok alapján az egyén költségének becslését végző fejkvóta formulát ($c(\cdot)$) alkalmasan megválasszuk.

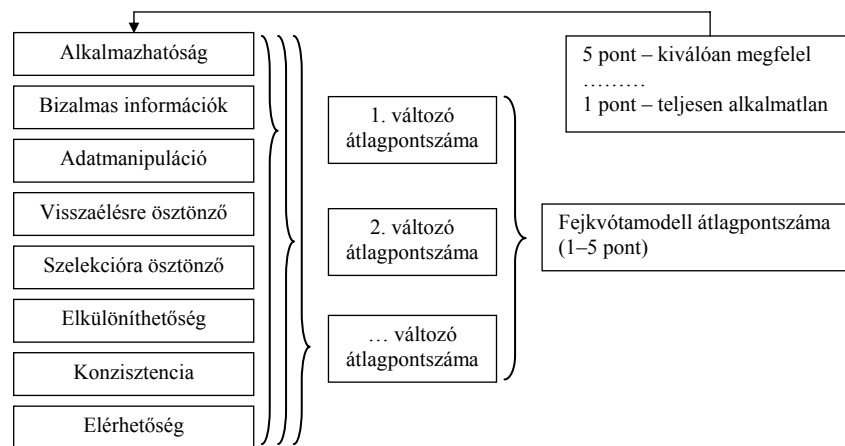
Az előbbieket alapján első látásra a fejkvótaszámítás számos hasonlóságot mutat a biztosítótársaságoknál végzett – a kockázati díjszabás készítésére alkalmazott – elemzésekkel, hiszen adott kockázati tényezők alapján kell bizonyos emberek várható kiadásait becsülni. A látszat ellenére azonban lényeges eltérések vannak. Az egészségügyi forrásallokáció esetén ugyanis nemcsak a kockázatok becslése és hatékony szétterítése a cél, hanem olyan egyéb szempontok figyelembe vétele is, mint az emberi szükségletek alapján történő igazságos elosztás, vagy a finanszírozott szervezetek hatékony működésre ösztönzése (Rice–Smith [2002], Nagy–Dózsa–Boncz [2004]). Ennek köszönhetően a formula készítése túlmutat a kockázatot leíró változók megtalálásán és képletbe illesztésén; valamint számos kiegészítő eljárás együttes alkalmazását feltételezi, melyek elvégzése után készülhet csak el a végleges kockázatkiigazítási formula.

1.2. Modellek építése

A kockázatkiigazító modellek készítésének első szakasza a kockázatbecslés, amely során az egyén várható egészségügyi kockázatát meghatározó változók kiválasztása történik; a második szakasz – a tényleges kockázatkiigazítás – a kockázatbecslésnél kiválasztott változók szelekciója alapján történő fejkvóta kialakítását foglalja magában. A két szakasz közötti fontos különbség, hogy a becslésnél feltárt statisztikai szempontból „eminens” változók nem minden esetben egyeznek meg a kiigazításhoz használt változókkal. Ugyanis a statisztikai teljesítmény mellett szá-

mos más kritériumnak is meg kell felelni; például annak, hogy a fejkvótaához rendelt változókat a finanszírozott szervezetek ne tudják a betegek, illetve a finanszírozó rovására befolyásolni; vagy annak, hogy csak az emberek valódi szükségleteivel összefüggő (ún. legitim) változók kerüljenek a fejkvótamodellbe (Nagy [2006]). Sok formulaválasztási szempont nehezen számszerűsíthető. Ilyen például a változók minősége, alkalmazhatóságuk, hatásuk a szabályozásra, az ösztönzésre, az igazságos elosztásra, illetve a változók bevezetésével járó adatmanipuláció veszélye. Ezek a szempontok országoként igen eltérők és nehezen általánosíthatók. Értékelésükhöz Rice és Smith [2001] készítettek egy nyolcelemű szempontrendszert, amelyet az 1. ábra tartalmaz. Ezeknek a szempontoknak a mérlegelése után könnyen elfordulhat, hogy a formula elkészítésekor nem a statisztikai szempontokból legjobban magyarázó modellt kell kiválasztani. Ez a sajátosság erősíti a modellépítés számos visszacsatolással működő iteratív jellegét, amelyben a – változókat és az allokációs egyenletet meghatározó – döntések jelentős része matematikai-statisztikai vizsgálatokra támaszkodik, de egyéb szempontokat is gondosan mérlegel.

1. ábra. A modellváltozók értékelése Rice és Smith szempontjai alapján



Forrás: Rice–Smith [2001].

1.3. A fejkvóta változói

A fejkvóta vektorához rendelhető tulajdonságok, azaz a fejkvóta finomításához használt változók köre igen széles. Extrém esetben léteznek rendkívül egyszerű kockázatiigazítási modellek, akár olyanok is, amelyekben minden emberhez ugyanazt

az összeget rendelik (például Spanyolországban), de ez egyre ritkább. A leggyakoribb változók a demográfiai tényezők (kor és nem), ezek mellé a fokozatos fejlesztés során új elemeket rendelnek. Alapvetően két változócsoporthoz különböztethető meg: az egészségi állapottal közvetlen kapcsolatban álló *egészségváltozók*, amelyeket általában diagnosztikus, gyógyszer-felhasználási adatok alapján alakítottak ki; illetve a főként szociodemográfiai, foglalkoztatási, infrastrukturális, háztartási, iskolázottsági, mortalitási és egyéb adatokon alapuló *nem egészségváltozók*. Az egészségváltozók egy korábbi időszak gyógyszer- (lásd például *Lamers–Vliet* [2004], *Zhao et al.* [2005], *Sales et al.* [2003], *Fishman et al.* [2003], *Gilmer et al.* [2001] modelljeit), járó- (lásd például *Weiner et al.* [1996], *Welch* [2002] modelljeit) és fekvőbeteg-ellátási (lásd például *Ellis et al.* [1996], *Ash et al.* [2000], *Kronick et al.* [1996] modelljeit) információi alapján képeznek olyan – általában a krónikus betegségben szenvedőkre jellemző – kockázati csoportokat, amelyek esetében a költségek jól becsülhetők. Ezek a modellek az egészségügyi adatok részletes feldolgozását végző szoftverek segítségével működnek, és rendkívül jól használhatók a várható egészségügyi költségek becsülésére. A nem egészségváltozók ugyan általában könnyebben és gyorsabban elérhetők és előállíthatók, mint az egészségváltozók, mivel rutinszerűen előállított országos adatbázisokban gyakran megtalálhatók. Viszont lényeges hátrányuk, hogy egyénekre bontva ritkán állnak rendelkezésre, és kapcsolatuk az egészségügyi kiadásokkal nem mindig egyértelmű. Ezért használatuk ugyan kézenfekvő és gyakori, de kevésbé ajánlott.

1.4. A függvényforma

A korábban leírt $c(\cdot)$ függvény megválasztása – azaz a várható költségek modellezése – abból a jól ismert feltevésből indul ki, hogy az egészségügyi költségek eloszlása erős ferdeségét mutat. Ezt még jobban eltorzítja (elnyújtja) a populáció egészséges része. Magyarországon például az egészségügyi kiadások 64 százalékát a betegek 10 százaléka használja fel, miközben az emberek 11 százalékánál gyakorlatilag semmilyen költség nem merül fel (*Nagy–Dózsa* [2002]). Az ilyen kis létszámú, de nagy költségű betegpopuláció modellezésére használt klasszikus becslőmodell két lépésből áll (*Duan* [1983]). A modell az első lépésben meghatározza, hogy a populáció egy adott eleménél előfordult-e a betegség (bináris változó, amelynek becslése probit/logit valószínűségi modellekkel becsülhető), majd feltéve, hogy előfordult, elvégzi az előfordulás után jelentkező költségek becsülését (szigorúan pozitív változó, amelynek becslésére általában a lineáris regresszió eszköztárából válogatnak, sokszor logaritmikus transzformációt alkalmazva). (Lásd erről *Jones* [2000], *Manning et al.* [2003], *Manning et al.* [2005], *Buntin–Zaslavsky* [2004].) Több elemző rámutat arra, hogy megfelelően nagy mintanagyságnál már

az egész populáció költségeire illesztett lineáris modellek, sőt akár az egyszerű csoportosítások is legalább olyan jól teljesítenek, mint a klasszikus modell. (Lásd erről *Van de Ven–Ellis* [2000], *Mullahay* [1998], *Ellis–Azzone* [1998], *Shen–Ellis* [2002].) Mivel napjainkban egyre nagyobb betegpopulációkról állnak rendelkezésre adatok, ezért a klasszikus modell használatát a gyakorlatban egyre inkább az említett egyszerűbb modellek váltják fel.

1.5. A modellek teljesítményének mérése

A fejkvótás modellek teljesítményét a szakirodalomban kedvező statisztikai tulajdonságai miatt leggyakrabban az R^2 -mutatóval mérik, amely azt mutatja meg, hogy a modell változói a valós költségek varianciájának hány százalékát képesek magyarázni. Az elvárható maximumérték (varianciahányad), amit az egyén várható összes egészségügyi költségeiből egy keresztmetszeti prospektív modell képes lehet megbecsülni 20-25 százalék körül van (*Newhouse* [1996], *van Vliet* [1992]). A gyakorlat is ezt támasztja alá, kevés modell ér el 20 százalék feletti R^2 -értéket. Ez egyáltalán nem meglepő, hiszen a fennmaradó 70-80 százalék nagy részét – például egy lábtörést, rákos daganat diagnosztizálását, vagy terhesség kialakulását – legtöbb esetben nem lehet előre jelezni.¹ A modellek statisztikai teljesítménye jelentősen változhat a vizsgált szolgáltatások, az érintett populáció, a becült időszak hossza, és a felhasznált változók minőségének függvényében. Bizonyos kockázati alcsoportokat külön-külön vizsgálva nagy különbségeket találhatunk a magyarázóerőben, például idősebb korcsoportoknál általában jobb a modellek magyarázóereje (*Newhouse* [1998]) (ez leginkább annak köszönhető, hogy az idősebbeknél a költségek magasabbak és gyakoribbak). Jelentős különbséget találunk akkor is, ha csak kórházi ellátásra ($R^2=0,05$), csak járóbeteg-ellátási költségekre ($R^2=0,25$) (*Newhouse et al.* [1989]), vagy csak gyógyszerköltségekre vonatkoztatva készítünk becsléseket ($R^2=40\%$) (*Andersson et al.* [2000]).

Az R^2 komoly hátránya, hogy igen érzékeny az ún. kiugró értékekre, bizonyos esetben például 10 000 darab 10 000 forintnyi – azaz nagy számú viszonylag alacsony – eltérést ugyanúgy értékel, mint 1 darab magas, 1 000 000 forintos eltérést (*Van de Ven–Ellis* [2000]). Így akár néhány rendkívül kiugró érték, adatjelentési hibák vagy visszaélések erőteljesen torzíthatják az eredmény értékelését. Az R^2 hibái a kiugró esetek kiszűrésével kezelhetők, noha ezzel a regresszió elveszíti a „zéroátlag” tulajdonságát, azaz a becült költségek összege a levágott értékek összegével kevesebb költséget mutat. Az R^2 hátrányos tulajdonságai miatt gyakori más további mutatók használata. Ilyen az átlagos abszolút eltérés (Mean Absolute Deviation – MAD),

¹ A modellek általában egyéves időtartamra vonatkozó becsléseket végeznek.

amely a becslt és a valódi értékek közötti abszolút eltérést vizsgálja. A MAD az R^2 -tel szemben sokkal kevésbé torzul a kiugróan magas hibák esetén (hiszen ezeket nem emeljük négyzetre), másrészt a megfigyelt és a modell által predikált eltérések nagyságrendjét is tükrözi. Hátránya viszont a ritkább gyakorlati alkalmazás, így nehéz összehasonlítani, hogy milyen standard érték vagy változás tekinthető jónak vagy jobbnak, mint a másik. Az ide vonatkozó szakirodalom említ még néhány mérőszámot, amelyek jellegüket tekintve a MAD-hoz nagyon hasonlóak (*Behrend et al.* [2007], *Ash et al.* [2000]).

A modellek teljesítményének összehasonlítására gyakran használt kiegészítő módszer bizonyos mintavételi eljárások alkalmazása is. Ezeknek kétféle szerepük van: egyrészt a becslések több mintára való elvégzésével információt szolgáltatnak a finanszírozott szervezetek várható költségeinek lehetséges ingadozásáról; másrészt a több évre vonatkozó minták kialakításával lehetőséget adnak nemcsak a statisztikai teljesítmény, de a modellek hosszabb időszakokra vonatkozó viselkedésének elemzésére.

A különféle modellek teljesítményének összehasonlítása igen korlátos, nagyban függ a fejkvótával lefedett (becsülni kívánt) ellátások körétől, a használt adatok jellegétől (individuális, vagy aggregált) és a teljesítmény mérésére használt mérőszámoktól. Nem is biztos, hogy az összehasonlítás feltétlenül szükséges, hiszen minden modellnek a saját alkalmazási környezetében kell elősegítenie a hatékony és méltányos forráselosztást. Fontos azt is látni, hogy a statisztikai teljesítmény növelése nem mindenáron kívánatos. Sokszor fordított összefüggés figyelhető meg a változók bevonásával növekvő statisztikai teljesítmény és ugyanezen változók bevonása miatti nem kívánatos egészségügyi ellátási ösztönzők erősödése között; ezért a statisztikai eredményeket, mindig az adott körülményekkel együtt kell értelmezni, értékelni.

1.6. Modellválasztási szempontok

Ahogy a változók kiválasztása, úgy a modellek közötti választás is számos preferenciát tükröz a statisztikai teljesítmény mellett. Ilyen például az ellátások közötti elosztás hatékonyságának növelése, az igazságos és méltányos elosztási rendszer erősítése, a megfelelő szolgáltató-ösztönzési struktúra kialakítása, vagy éppen a nem legitim változók kiküszöbölése. A társadalom értékrendje alapján különféle választási szempontrendszer alakítható ki, melyek mögött eltérő statisztikai modellválasztási helyzetek találhatók. Annak ellenére, hogy a társadalmi értékválasztás nem jelöl ki egyetlen helyes modellépítési utat, általában előnyökkel és hátrányokkal kevert megoldások felé orientálja a fejkvóta készítőit. Ezeket a modellválasztási szempontokat *Van de Ven* és *Ellis* [2000] három csoportba rendezi. A fejkvóta: 1.

mint ösztönző hogyan működik, 2. mennyire igazságos, és 3. a gyakorlatban hogyan alkalmazható.

1. A cél olyan ösztönzőket keresni és bevezetni, amelyek a megfelelő irányba mozdítják az ellátások minőségét és mennyiségét. Ezért csak megfelelő „védelem” mellett szabad alkalmazni olyan tényezőket, amelyek javítják ugyan a formula statisztikai teljesítményét, de az egészségügyi szolgáltatók viselkedését rossz irányba befolyásolják.

2. Az igazságos formula kialakításának a felek számára elfogadható, szolidaritási és más igazságossági elveket is értékelő elosztórendszer kialakítását kell tükröznie. Objektív igazságossági elveket rendszerszerűen kidolgozni és számszerűsíteni rendkívül nehéz, kevés országban tettek erre kísérletet. Ennek ellenére minden egészségügyi rendszer egyik legfontosabb eleme ez a kritérium.

3. Az alkalmazhatósági feltétel a gyakorlati megvalósítás részleteit vizsgálja. A könnyen, egyszerűen hozzáférhető, egyébként is rutinszerűen gyűjtött adatok a fejkvótamodellek fejlesztésének alapfeltételei. A változók kiválasztása után a hangsúly az adatbázisok validitásának növelésén, illetve a tudományos igényű modellekhez szükséges adatok elérési feltételeinek megteremtésén van. A végső alkalmazás során a technikai tökéletesség mellett az egyszerűség iránti igény is gyakran fontos szempontként fogalmazódik meg (*Smith et al.* [2001]).

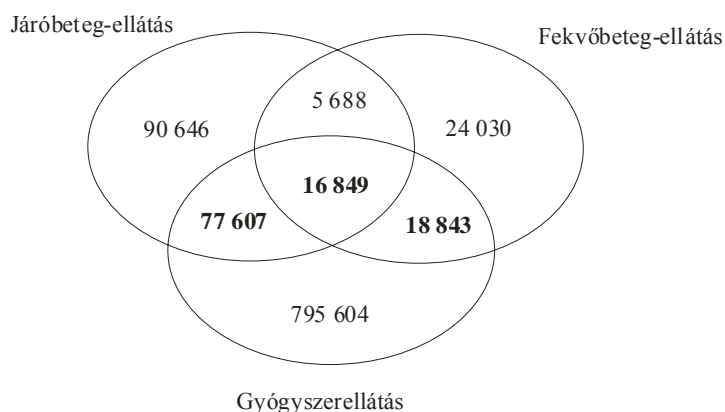
2. Empirikus elemzés: a COPD-betegek fejkvótájának kialakítása

A következő példaszámítás azt illusztrálja, hogy a tanulmány elméleti részében felvázolt fejkvóta-készítési szempontok alapján egy adott alkalmazási környezetben milyen megoldási lehetőségek kínálkoznak. Az elemzés egy betegcsoport fejkvótájának kialakításán keresztül mutatja be a kockázatbecslés és a kockázatkiszármazás folyamatát. A krónikus obstruktív tüdőbetegségben (Chronic Obstructive Pulmonary Disease – COPD) szenvedők fejkvótájának kialakítását vizsgáljuk rendelkezésre álló magyar egészségügyi adatok felhasználásával. A COPD-betegek adataira alapozva számos egészségi adatokat használó fejkvótamodell készült már, mivel betegek köre jól definiált, a diagnózis és a gyógyszerfogyasztási mutatók viszonylag könnyen gyűjthetők, és a betegség jellegéből adódóan a várható kockázat is jól becsülhető.

2.1. Adatok

A Magyarországon 2002. augusztus – 2003. július időszak alatt egészségügyi fekvőbeteg-, járóbeteg- vagy gyógyszerellátásban megjelenő COPD-betegek vizsgálatára került sor.² Az Országos Egészségbiztosítási Pénztár (OEP) adatállományának betegsintű adatait válogattuk le, az adatvédelmi szabályok betartásával, személyes azonosításra alkalmatlan módon. A COPD-populáció a betegséghez rendelhető nemzetközi betegség- (Betegségek Nemzetközi Osztályozása – BNO) és gyógyszer-osztályozási (Anatomical Therapeutic Chemical classification – ATC) kódok alapján került kiválasztásra.³ Az első válogatás során talált betegek körét szűkíteni kellett, mivel az egyes ellátási formáknál megjelenő kódolási hibákból, illetve a hasonló betegségekkel történő átfedésekből eredő nem valódi COPD-betegek téves kiválasztása így vált elkerülhetővé. Ezért az eredeti állományból azokat a betegeket választottuk ki, akik a vizsgált időszakban legalább egyszer COPD kezelésére jellemzően alkalmas gyógyszert váltottak ki és járó-, vagy fekvőbetegellátásban az adott diagnózissal megjelentek. A válogatás végeredményét a 2. ábra vastagított számai mutatják, a kiválasztott betegek leíró statisztikáit pedig a 1. táblázat tartalmazza. A táblázatból látható, hogy a betegek költségei igen nagy tartományban szóródnak (átlag: 58 193 forint, minimum: 310 forint, maximum: 2,8 millió forint), ami a költségek szempontjából igen heterogén betegcsoportra utal.

2. ábra. A járó-, fekvő- és gyógyszerellátások alapján kiválogatott COPD-betegek száma



² Néhány COPD-kezeléssel kapcsolatos ellátás vizsgálatára adathiány miatt nem volt lehetőség: ilyen a gondozóintézeti ellátás, a házi szakápolás és a gyógyászati-segédeszköz ellátás.

³ Az ATC és a BNO osztályozási rendszer a számos egészségügyi adatokat használó fejkvótamodell kiindulópontja. A kódok listáját *dr. Szilasi Mária*, a Debreceni Egyetem Tüdőgyógyászati Klinika docensének közreműködésével alakítottuk ki. Az adatlekérdezési stratégiát az OEP munkatársai, *Brandtmüller Ágnes* és *Reszegi Csilla* készítették.

1. táblázat

A kiválasztásra került COPD-betegpopuláció leíró statisztikája

Változó	Ellátási forma			
	Gyógyszer	Fekvőbeteg	Járóbeteg	Összes ellátás
	Költség (forint)**			
Átlag*	9 924	45 795	2 474	58 193
Átlag	9 924	158 747	2 968	58 193
Medián	3 960	120 938	1 852	12 600
Szórás	16 964	156 382	3 602	112 017
Minimum	175	2 142	7	310
Maximum	311 133	2 783 119	192 384	2 783 119
	Átlag**			
Életkor	53,5	55,2	53,1	53,5
Férfiak (százalék)	49,0	55,0	51,0	49,0
Létszám	113 299	35 692	94 456	113 299

* A összes betegre (113 299) vetítve.

** A vizsgált ellátási típusnál megjelent betegekre vonatkoztatva.

A fejkvóta kialakításához szükséges magyarázóváltozók elkészítéséhez az adatbázisban található demográfiai információkat (kor és nem), illetve az egészségügyi szolgáltatóknál jelentett adatokat használtuk. A szolgáltatóknál jelentett adatok három csoportját különböztettük meg.

1. *Teljesítményváltozók*, amelyek az egészségügyi szolgáltatóknak megállapított finanszírozási összeg (elemzésünkben ezek a költségek) kiszámításához használatos mennyiségi információkat hordozzák.
2. *Kombinált változók*, amelyek az ellátások mennyiségi paramétere mellett egyéb tartalmat is hordoznak (például az ellátás típusa).
3. *Csoportváltozók*, amelyek a teljesítményváltozók különféle leképezéseit jelentik kettő vagy három halmazra.

Összesen 35 magyarázóváltozó kialakítására került sor, melyeket a 2. táblázat tartalmaz. A 2. táblázatból látható, hogy egy COPD-beteg az év során átlagosan hétszer jelent meg járóbeteg-ellátási intézetekben (*jeset*), 110 napra elegendő gyógyszert vett ki (*dotsum*) és a betegek 10 százaléka több mint 12 napig volt kórházban.

2.2. A modellépítés és -értékelés módszere

A modellek építésének első fázisa a kockázatbecslés volt, amely során olyan regressziós modelleket alakítottunk ki, amelyekben alulról felfele építkezéssel (ún. stepwise regresszióval) egyre több változóval, egyre bonyolultabb modelleket becsültünk. A kiinduló formulákban kizárólag konstansok szerepeltek, melyekhez lépésenként egy-egy újabb változót vettünk hozzá, mégpedig olyan módon, hogy minden lépésben a legnagyobb magyarázóerejű modell változójával bővítettünk. Az algoritmus során az új változó modellbe kerülésének határa az 1 százalékos szignifikanciaszint volt. A hasonló üzenetet hordozó változók multikollinearitását segédregresszióval vizsgáltuk: a kiszemelt, új változót a régiekkel közelítve meghatározott R^2 felett ($R^2 > 70\%; 40\%; 10\%$) az új változót multikollineárisnak tekintettük és elvetettük, majd helyette újat kerestünk. Azért, hogy a modellek érzékenységét változócsoportokra is demonstrálni lehessen az előző iteratív modellépítési folyamatot:

- az összes változón (*teljes modellek*),
- a csak dummy és csoportváltozókon (*dummy modellek*),
- a fekvőbetegellátási dummy változón (*fekvő* – lásd a 2. táblázatot) és az összes gyógyszer- és járóbeteg-ellátási változón (*fekvő+többi modellek*), illetve
- a járó- és gyógyszerváltozókon (*j_gy modellek*)

külön-külön is elvégeztük. Emellett azt is megvizsgáltuk, hogy az egyes ellátási típusoknál jelentkező költségeket más ellátások változóival mennyire lehet becsülni. Az így felépített „mintamodellek” statisztikai teljesítménye benchmark-ként szolgált a végső fejkvóta felépítéséhez.

A modellépítés második fázisában (kockázatkiigazítás) a kockázatbecslés eredményeit felhasználva a potenciális változókat egyenként, illetve különböző kombinációkban „manuálisan” vizsgáltuk. Ezeknek a végső modelleknek kialakításában már nem csak a statisztikai szempontok kaptak szerepet, hanem a változók minősége, a modellek forrásallokációs hatása és olyan elbírálási szempontok, mint a modellek nagysága, várható hatásuk a finanszírozott szervezetek viselkedésére, illetve a változókkal hatékonyan elkülöníthető kockázati csoportok mérete.

A modelleket három mérőszám alapján értékeltük:

1. a modellek változóinak *Rice–Smith* [2001] -féle pontértéke (lásd az 1. ábrát);
2. magyarázóerő (R^2);
3. és a prediktív erő (MAD), azaz a valós és a becsült költségek átlagos abszolút eltérése.

Az iteratív modellépítés értékelése (magyarázóerő-, multikollinearitás- és szignifikanciavizsgálat) és a statisztikai szempontból elfogadható modellek kiválasztása a rendelkezésre álló teljes COPD-populáció elemzésével történt. A változók pontértékeit *Rice–Smith* [2001] szempontrendszere alapján (lásd az 1. ábrát) 1-5-ig terjedő skálán értékeltük, és minden modellnél megvizsgáltuk. A prediktív erő vizsgálatához az ún. bootstrap eljárást alkalmaztuk, melynek során a teljes populációt visszatevés nélküli véletlen mintavétellel két részre bontottuk, és az adatok egyik részén (az adatok 70 százaléka; *becslő minta*) a kockázatbecsléssel kiválasztott modellek paramétereit megbecsültük, majd a másik részén (az adatok 30 százaléka; *értékelő minta*) ugyanezeknek a becsült modelleknek a prediktív erejét mértük. Ezt az eljárást 100-szor megismételtük és a 100 eredmény átlagából számoltuk ki a végső értékeket. A prediktív erőt virtuálisan kialakított ellátásszervezők szintjén mértük.⁴

2.3. A változók Rice–Smith-pontszámai és az iteratív modellépítés eredményei

A magyarázóváltozók átlagos értékét és a *Rice–Smith* [2001] -féle pontozás eredményeit a 2. táblázat tartalmazza. A teljesítményváltozók értéke a legalacsonyabb, ennek fő oka, hogy ezek igen érzékenyek az adatmanipulációra, és visszaélésekre ösztönöznek az ellátásban. Mivel a kombinált, a dummy- és a csoportváltozók kevésbé érzékenyek az adatmanipulációra és kisebb a negatív ösztönző hatásuk, kedvezőbb értékelés kapnak. A demográfiai változók értékelése a legjobb, szinte minden kategóriában maximális pontot értek el. Ennek megfelelően a modellek későbbi „manuális” építéskor a Rice–Smith-kritériumok alapján a demográfiai és a magasabb értékkel bíró minőség-, dummy és csoportváltozók előnyt élvezhetnek.

A különféle változócsoportokon elvégzett iteratív modellépítés eredményeit a 2. táblázat tartalmazza. A szisztematikus szelekció ellenére láthatóan sok változó maradt a modellekben: még a legszigorúbb modellépítési feltételek mellett is 35-ből legalább 4, de jellemzően 8-10 változó szignifikánsnak bizonyult. A 3. táblázatból az is látszik, hogy a fekvőbeteg-ellátásból nyert változóknak (*fekvo*, *sumapnap*) minden releváns modellben kulcsszerep jut: elsőként kerülnek felvételre és nélkülük a költségek varianciájának maximum 15-16 százalékát lehet megmagyarázni (lásd *j_gy modellek*). Az egyetlen fekvőbeteg-ellátási változót használó modellek (*fekvő+többi modellek*) azt is jól mutatják, hogy a kórházba kerülés tényének ismerete – kiegészít-

⁴ Az ellátásszervezőnkénti bontás nem valódi, hanem random módon létrehozott 8 hasonló létszámú szervezet kialakítását jelentette. A prediktív erő ellátásszervezőnkénti elemzésénél a szervező létszámának gyökével súlyoztuk az abszolút eltérések átlagát. A „mintamodellek” építése az EVIEWS 4.1 adatelemző szoftverrel, a bootstrap eljárás R 2.5.1 statisztikai programcsomaggal, minden más statisztikai elemzés az SPSS 14 statisztikai programmal készült.

ve a többi ellátási formából nyert információkkal – önmagában is igen magas (körülbelül 40 százalékos) magyarázóerőt produkál. Ugyanakkor az egyes fekvőbeteg-ellátási változók egymáshoz hasonló információkat hordoznak: 10 százalékos multikollinearitási feltétel mellett már csak egyetlen fekvőbeteg-ellátási változó marad minden vizsgált modellben.

2. táblázat

A COPD-modellek építéséhez használt magyarázóváltozók

Typus	Név	Meghatározás	Változók értékének átlaga	Rice-Smith pontszám
teljesítményváltozók	jeset	Éves esetszám a járóbeteg-ellátásban	7,2	3,125
	feset	Éves esetszám a fekvőbeteg-ellátásban	0,6	3,125
	gyeset	Patikai vásárlások esetszáma	5,5	3,125
kombinált változók	countbno	BNO típusok száma az év során	1,0	3,375
	sumapnap	Összes évi fekvőbeteg-ellátásban töltött napok száma	4,3	3,250
	countvdr	HBCS típusok száma az év során	0,4	3,250
	sumdot	Minden gyógyszerre összesített átlagos éves fogyasztás napi dózisban	109,8	3,375
	atctip	Szedett gyógyszerek ATC típusainak száma	1,9	3,250
demográfiai változók	nem	férfi: 1; nő: 2	1,5	4,750
	kor	Adott évre megadott egész szám	53,5	4,750
dummy változók	jesetavg	Az átlagosnál többször fordult meg járóbeteg-ellátásban (1-igen; 0-nem)	0,2	3,625
	jeset2 / jese6 / jese7	Több mint 2/6/7 alkalommal volt járóbeteg-ellátáson (1-igen; 0-nem)	0,7 / 0,4 / 0,3	3,375
	jeset12	Több mint 12 alkalommal volt járóbeteg-ellátáson (1-igen; 0-nem)	0,2	3,500
	jeset52 / jese150	Több mint 52/150 alkalommal volt járóbeteg-ellátáson (1-igen; 0-nem)	0 / 0	3,625
	apnapavg	Az átlagosnál több időt töltött kórházban (1-igen; 0-nem)	0,1	3,625
	fekv	Részesült fekvőbeteg-ellátásban (1-igen; 0-nem)	0,3	3,375
	apnap12	Több mint 12 napig volt kórházban (1-igen; 0-nem)	0,1	3,500
	dotavg	Az átlagosnál több volt az éves DOT fogyasztása (1-igen; 0-nem)	0,3	3,625
	dot50 / dot100 / dot120 / dot180	Szedett több mint 50/100/150/180 napnyi gyógyszert (1-igen; 0-nem)	0,6 / 0,3 / 0,3 / 0	3,625
	gyeset6	Több mint 6 alkalommal vásárolt gyógyszert (1-igen; 0-nem)	0,3	3,375
	Gyeseavg	Átlagosnál többször vásárolt gyógyszert (1-igen; 0-nem)	0,3	3,500
	csoportváltozók	copdcs01	2 - összes fekvő beteg; 1 - sumdot>180 nap; 0 - sumdot<180 nap	0,6
copdcs02		2 - összes fekvő beteg; 1 - sumdot>100 nap; 0 - sumdot<100 nap	0,8	3,500
copdcs03		2 - összes fekvő beteg; 1 - sumdot>50 nap; 0 - sumdot<50 nap	1,0	3,500
copdcs04		2 - összes fekvő beteg; 1 - sumdot>120 nap; 0 - sumdot<120 nap	0,8	3,500
copdcs05		2 - összes fekvő beteg; 1 - jese6>2; 0 - jese6<2	1,2	3,500
copdcs06		2 - összes fekvő beteg; 1 - jese6>6; 0 - jese6<6	0,9	3,500
copdcs07		2 - összes fekvő beteg; 1 - jese6>12; 0 - jese6<12	0,7	3,500
copdcs08		2 - összes fekvő beteg; 1 - jese6>52; 0 - jese6<52	0,6	3,500

Az iteratív módon felépített modellek további fontos eredménye, hogy mindhárom ellátási forma adekvát információt tudott szolgáltatni a költségek becsléséhez. A járó- és gyógyszerellátási változók hozzájárulása a magyarázóerőhöz ugyan csak néhány további százalék, de ezek a változók statisztikailag szignifikánsak maradnak a

fekvőbeteg-ellátási változók mellett. Használatukat különösen indokoltá teszi, hogy a betegek jelentős részének (68 százalék; lásd az 1. táblázatot) költségeit kizárólag ezek a változók képesek tovább differenciálni. Az is jól látható, hogy noha általában sok szignifikáns magyarázóváltozó marad a modellekben, de viszonylag kevés változó (3-4) alkalmazásával is lehet jó magyarázóerejű becslőmodellt készíteni.

3. táblázat

A mintamodellek eredményei különböző multikollinearitási feltételek mellett

Multikollinearitás vizsgálata	Vizsgált változócsoporthoz	R ² (százalék)	Szignifikancia változók száma	A végső modell változói (felvételük sorrendjében)
Nincs	teljes	59,6	13	sumapnap, countvdr, gyeset, copdcso8, feset, kor, sumdot, apnap12, jeket, jeket2, copdcso5, dot50, fekvó
	dummy	47,8	15	fekvó, apnapavg, gyeset6, dot120, jeket52, copdcso4, kor, apnap12, copdcso8, jeket12, dot180, copdcso7, nem, jeket150, gyesavg
	fekvő+többi	41,7	12	fekvó, gyeset, jeket, sumdot, countbno, kor, jeket6, dot180, dot50, nem, jeket2, jeket12
	j_gy	15,9	13	sumdot, jeket2, jeket, gyeset, nem, atctip, jeket150, kor, dot180, jeket12, countbno, dot120, dot50
70 százalék	teljes	59,2	9	sumapnap, countvdr, gyeset, copdcso6, kor, sumdot, jeket6, jeket, dot50
	dummy	47,5	11	fekvó, apnapavg, gyeset6, dot120, jeket52, kor, jeket12, dot180, nem, jeket150, dot50
	fekvő+többi	41,7	12	fekvó, gyeset, jeket, sumdot, countbno, kor, jeket6, dot180, dot50, nem, jeket2, jeket12
	j_gy	15,9	13	sumdot, jeket2, jeket, gyeset, nem, atctip, jeket150, kor, dot180, jeket12, countbno, dot120, dot50
40 százalék	teljes	57,6	10	sumapnap, copdcso4, gyeset, kor, dot50, jeket52, jeket2, jeket12, nem, dot180
	dummy	47,5	11	fekvó, apnapavg, gyeset6, dot120, jeket52, kor, jeket12, dot180, nem, jeket150, dot50
	fekvő+többi	41,4	9	fekvó, gyeset, jeket, countbno, dot100, kor, jeket6, dot180, nem
	j_gy	15,5	8	sumdot, jeket2, jeket, gyeset6, nem, dot180, kor, jeket150
10 százalék	teljes	54,7	8	sumapnap, gyeset, jeket2, kor, jeket12, jeket52, nem, dot180
	dummy	39,8	7	fekvó, dot120, jeket52, jeket12, dot180, kor, nem
	fekvő+többi	41,2	4	fekvó, gyeset, jeket, dot180
	j_gy	14,8	6	sumdot, jeket2, jeket12, jeket52, nem, kor

A különféle ellátások költségeit más ellátások változóival csak kis mértékben lehetett magyarázni. A járóbeteg-ellátási költségek varianciájának 4-5 százalékát, a kórházi költségek 11 százalékát, illetve a gyógyszerköltségek 3-4 százalékát lehetett a többi ellátás változóival magyarázni, amely messze elmarad a saját ellátási mutatókat is használó modellek teljesítményétől. (Lásd a 3. táblázatot.)

Összegezve, az iteratív modellépítési eredmények alapján a COPD-modellek építésének objektív feltételül szabható, hogy legalább fekvőbeteg-ellátási információt,

és ha van rá lehetőség, akkor a többi ellátási formából nyert információt alkalmazzanak. Ennek megfelelően a manuális modellépítéskor olyan modellek kialakítását érdemes vizsgálni, amelyek 1. korlátozott mértékben, de mindhárom ellátási formából használnak adatokat, 2. az iteratív statisztikai módszerekkel kialakított modellekhez közelítő teljesítményt képesek produkálni 3. és mindezt jó minőségű változók segítségével teszik.

2.4. A manuális modellépítés eredményei

Az iteratív modellek iránymutatásainak megfelelően a manuális modellépítéshez első lépésként a három ellátási forma legjobb változóit válogattuk ki a *Rice–Smith* [2001] -féle értékelés alapján. (Lásd az 1. ábrát.) A fekvőbeteg-ellátásnál a *fekvo*, *apnapavg*, *apnap12*; a gyógyszerellátásnál a *dot120*; a járóbeteg-ellátásnál a *jeset52* és a *jesetavg* változók bizonyultak a legalkalmasabbnak. E hat változó kombinációit vizsgálva a modellépítési kritériumoknak leginkább megfelelő modelleket kerestük; ezek prediktív erejét (MAD) a bootstrap eljárással teszteltük. Az eredményeket a 4. táblázat tartalmazza. A táblázatban az 1. modell csak egyetlen változót használ, a 2. és 3. modell a fekvő-betegellátási változó (*fekvo*) mellé egy járó- (*jeset2*) vagy gyógyszerellátási (*dot120*) változót rendel, a 4. modell a 3 ellátási forma változóinak együttes használatával kialakított legjobb kombinációt tartalmazza. A demográfiai (5.) modell ismertsége, egyszerűsége és magas *Rice–Smith* [2001] -féle pontértékei miatt került a 4. táblázatban bemutatásra.

A 3. és 4. táblázatokból látható, hogy az iteratív modellek eredményeihez (3. táblázat) az 1-4 modellek (4. táblázat) statisztikai teljesítményei igen jól közelítenek. A változók számának csökkentése ellenére 40 százalék körül marad a magyarázóerő. A 4. táblázatból az is látható, hogy a magasabb R^2 -nek ára van: a 4. modellnek már 16 különböző cellát kell képeznie a 7-9 százalékkal magasabb magyarázóerő eléréséhez; eközben azonban a cellák egynegyedébe nagyon kevés beteg sorolható, és számos kategória nagyon hasonló nagyságú fejkvótát produkál – a 110 004 forintos és 156 370 forintos fejkvóták között 6 darab további kategória található. Az is látható, hogy a MAD az R^2 -hez hasonlóan értékeli a modelleket: a legkisebb MAD-dal a 4. modell rendelkezik, amely a legmagasabb R^2 -et produkálja. A változók *Rice–Smith* [2001] -féle értékelése minimális eltéréseket produkál; a 4. modell változóinak átlaga – elsősorban a kisebb manipulációs veszély és a gyengébb negatív ösztönző hatás miatt – árnyalatnyival jobb eredményeket mutat, de az 1-4 modellek között lényeges különbséget nehéz találni. Összességében a 4. modell az 1-3 modelleknél valamivel jobb teljesítményeket produkál, de ehhez a változók számának növelésére volt szükség. A demográfiai (5.) modell gyenge statisztikai teljesítménye miatt láthatóan nem alkalmas a COPD-betegek fejkvótájának finomítására.

4. táblázat

A kockázatküigazításhoz kiválasztott modellek összehasonlítása

Modell	Változók	Koefficiens	Rice–Smith pontszám	R ² (százalék)	MAD (százalék)	Fejkvóták értéke* (forint)	Betegszám a fejkvóta csoportokban
1. (fekvo_)	(konstans)	11 947	3,38	37,10	1,87	11 947	77 607
	fekvo	146 800				158 747	35 692
2. (fekvo_jeset2)	(konstans)	–922	3,38	37,30	1,85	13 722	68 197
	fekvo	153 355				152 433	20 303
	jeset2	14 645				167 077	15 389
						–**	9 409
3. (fekvo_dot120)	(konstans)	2 957	3,50	39,10	1,83	2 957	58 109
	fekvo	14 2923				145 880	22 858
	dot120	35 783				38 740	19 497
						181 663	12 834
4. (szelekt_jesetavg)	(konstans)	2 886	3,56	46,80	1,64	2 886	44 391
	fekvo	107 119				110 004	14 940
	apnapavg	117 908				11 897	13 716
	dot120	26 565				29 451	12 706
	jesetavg	9 011				38 462	6 789
						136 569	5 216
						227 913	4 545
						254 477	3 528
						119 016	2 236
						263 489	2 083
		145 580	2 007				
		236 924	1 137				
		147 359	2				
		120 794	1				
		129 806	1				
		156 370	0				
5. (demográfiai)	(konstans)	64 086	4,75	0,60	2,36		
	Nem	–12 814					
	Kor	247					

* A fejkvóták a modellek koefficienseinek értékei alapján kerültek kiszámításra.

** Negatív koefficiens esetén a fejkvóta 0 értéket vesz fel.

Megjegyzés. A 100*2 db fejkvóta cella helyhiány miatt nem kerül bemutatásra.

2.5. Választás a modellek között

A modellépítési vizsgálatok megmutatták, hogy a COPD-betegcsoport számára 3-4 változó segítségével igen jó statisztikai teljesítménnyel rendelkező súlyozott fejkvóta kialakítására van lehetőség. A végső elemzésekhez használt változók dummy

változóként kerültek meghatározásra. A kiválasztott négy modell (a demográfiai modellt elvetettük) a kórházi ellátási igénybevétel (*fekvo*), az eltöltött kórházi napok száma (*apnapavg*), a járóbeteg-ellátási igénybevétel (*jeset2*, *jesetavg*) és a gyógyszerfogyasztás (*dot120*) alapján rendezi 2-16 kategóriába a betegeket. Meglátásunk szerint a 4. táblázatban bemutatott 1-4 modellek közül bármelyik választható kockázatküszöböt meghaladóra, a döntést azonban kizárólag statisztikai kritériumok alapján nem lehet meghozni. Ezért ahogy korábban bemutattuk, a végső formula kialakításához fontos mérlegelni, hogy a modellek, mint *ösztönzők* hogyan működnek, mennyire *igazságosak*, és a gyakorlatban hogyan *alkalmazhatók*.

Ösztönző hatás. Mindegyik kiválasztott modell olyan információkra épül, amelyek az egészségügyi szolgáltatók által manipulálhatók, illetve ösztönözhetik őket visszaélésekre az ellátások nyújtása során. Ennek a veszélynek a mértéke nehezen becsülhető, de szinte minden egészségváltozókra épülő modell szembesül ezzel a problémával (*Lamers–Vliet* [2003], *Ellis* [2002]). A modellépítés során a dummy (és csoport-) változók kialakításával erőfeszítéseket tettünk ennek a veszélynek csökkentésére. Az, hogy valaki évente legalább 120 napig szedett, vagy nem szedett gyógyszert (*dot120*), meglátásunk szerint kevésbé manipulálható információ, mint a gyógyszerszedéssel eltöltött pontos napok száma (*sumdot*). Fontos látni azt is, hogy bizonyos határon túl a változók minősége között már nehéz különbséget tenni. Nehéz például eldönteni, hogy a 4. táblázatban látható 2. és 3. modell változóinak minősége között (*dot120* vagy *jeset2*) van-e számottevő különbség. Bármilyen döntés valószínűleg szubjektív mérlegelés tárgya lenne. (Az elemzés szerint a *dot120* változó kevésbé manipulálható és kevésbé ösztönöz visszaélésekre, mint a *jeset2*). Fontos azt is látni, hogy egy-egy dummy változó kiválasztása mellett orvos-szakmai érvek is szólhatnak. Például a COPD-betegek járóbeteg-ellátásban történő megjelenéseinek számát a kezelési irányelvekhez igazított dummy változó használata (például évente meghatározott számú vizit) helyes irányba is mozdíthatja.

Igazságosság. Magyarországon ismert az egészségügyi intézményi rendszer területi heterogenitása és az ebből adódó egyenlőtlen (igazságtalan) hozzáférés problémája (*Boncz et al.* [2006], *Szaszkó et al.* [2006], *Takács et al.* [2006]). Mindezek tudatában nehezen képzelhető el, hogy jelenleg a COPD-betegek az országban mindenhol pontosan ugyanolyan minőségű és mennyiségű ellátásban részesülnek. Ahol sűrűbb az intézményrendszer, ott valószínűleg jobb a betegek hozzáférése az ellátásokhoz, és gyakrabban járnak kórházakba és rendelőkbe; az ellátási gyakorlat is változhat intézményenként. Így elképzelhető, hogy a leírt modellek használatával két ugyanolyan súlyosságú COPD-beteg nem ugyanazt az egy főre eső összeget kapja, csupán azért mert egyikük könnyebben és gyakrabban jut el az orvosához. Az ilyen jellegű igazságtalanságokat területi információkat alkalmazó változók (például intézmény-lakóhely távolság) használatával lehet a fejkvóta formula készítése során figyelembe venni.

Alkalmazhatóság. A kiválasztott modellek alkalmazásának meglátásunk szerint nincs technikai akadálya. Könnyen reprodukálható, rutinszerűen gyűjtött és a viszonylag jó adatminőségű információk állnak rendelkezésre. Ez a körülmény megkönnyíti a modellek gyakorlati alkalmazását. Az egyszerűbb COPD-formulák (1. 2. 3. modellek) némi előnyt élveznek, főképp olyan esetekben, amikor sok betegségecsoportot tartalmazó fejkvótamodell alkotóelemeiként kerülhetnek bevezetésre. A formula kiválasztása előtt a döntéshozóknak mindenképpen érdemes mérlegelnie, hogy:

- milyen bonyolultságú fejkvótát szeretne készíteni;
- mely változók tűnnek az adott körülmények között jól alkalmazhatónak;
- melyik modellel, mennyire lehet elkerülni az adatmanipulációt;
- milyen akár az ellátás minőségének, hatékonyságának javítását elősegítő egészségügyi szabályozási elemeket kell rendelni a finanszírozási formula mellé;
- milyen orvos-szakmai szempontokat lehet és kell érvényesíteni a fejkvótában.

Ezen szempontok mérlegelése a matematikai-statisztikai ismeretek alkalmazásával párhuzamosan más diszciplínák közreműködését is igénylő fejkvóta-készítési gyakorlat kialakítását igényli, melyek részletes tárgyalása meghaladja ennek az dolgozatnak a kereteit.

3. Következtetések

Az egészségügyben használatos fejkvótaformulák készítésének nemzetközi tapasztalatai alapján számos lehetőség nyílik a statisztikai-matematikai módszerek alkalmazására a hazai fejkvóta-számítási gyakorlat kialakítása során. Ezek közül külön figyelmet érdemel a kockázatbecslési és kockázatkiigazítási műveletek szétválasztása, a statisztikai modellek építésének folyamata, a költségek becsléséhez használt függvényforma kiválasztása, a formula elkészítéséhez használt változók szelekciója és a fejkvótamodellek teljesítményének mérése. Az elemzésben bemutatott gyakorlati példa azt demonstrálja, hogy az alkalmazott modelltől függően akár egyetlen kockázati csoport (COPD-betegek) vizsgálata esetén is komoly eltérések adódhatnak az azonos betegségben szenvedő emberek várható egészségügyi ellátási költségeinek becslésekor. Az eltérések feltárása, illetve a bizonytalansági tényezők figyelembe vétele jelentősen befolyásolja a hatékony és méltányos forráselosztás sikerét. Indokolt

tehát, hogy a statisztikai módszerekkel kimutatott különbségek érvényesítése a finanszírozási formulában csak számos egyéb szempont mérlegelésével együtt történjen meg. Emiatt javasolt a fejkvóta számítást minden esetben széles szakmai, interdiszciplináris konszenzusra építeni, ugyanakkor statisztikai-matematikai módszerekkel megalapozni.

Irodalom

- ANDERSSON, P. A. – VARDE, E. – DIDERCHSEN F. [2000]: Modelling of resource allocation to health care authorities in Stockholm County. *Health Care Management Science*. 3. évf. 2. sz. 141–149. old.
- ASH, A. S. ET AL. [2000]: Using diagnoses to describe populations and predict costs. *Health Care Financing Review*. 21. évf. 3. sz. 7–28. old.
- BEHREND, C. ET AL. [2007]: Risk adjusted capitation payments: how well do principal inpatient diagnosis-based models work in the German situation? Results from a large dataset. *European Journal of Health Economics*. 8. évf. 8. sz. 31–39. old.
- BONCZ I. ET AL. [2006]: Az OEP krónikus fekvőbeteg szakellátási kassza igénybevétele. Területi egyenlőtlenségek I. *Kórház*. 12. évf. 7–8. sz. 37–43. old.
- BUNTIN, M. A. – ZASLAVSKY, A. M. [2004]: Too much ado about two-part models and transformation?: Comparing methods of modeling Medicare expenditures. *Journal of Health Economics*. 23. évf. 3. sz. 525–543. old.
- DUAN N [1983]: Smearing estimate: A nonparametric retransformation. method. *Journal of the American Statistical Association*. 78. évf. 383. sz. 605–690. old.
- ELLIS, R. [2002]: *Using fuzzy diagnoses for risk adjustment*.
<http://www.bu.edu/econ/workingpapers/papers/Randall%20P.%20Ellis/EllisFuzzyDiag5.pdf>.
Munkaanyag.
- ELLIS, R. – AZZONE, V. [1998]: *OLS, loglinear and two part models of health expenditure: what do the data tell us?* Munkaanyag.
- ELLIS, R. ET AL. [1996]: Diagnosis-based risk adjustment for Medicare capitation payments. *Health Care Financing Review*. 17. évf. 3. sz. 101–128. old.
- FISHMAN, P. A. ET AL. [2003]: Risk adjustment using automated ambulatory pharmacy data: the RxRisk model. *Medical Care*. 41. évf. 1. sz. 84–99. old.
- GILMER, T. ET AL. [2001]: The Medicaid Rx model: pharmacy-based risk adjustment for public programs. *Medical Care*. 39. sz. 11. sz. 1188–1202. old.
- JONES, A. [2000]: Health econometrics. In: *Culyer, A. J. – Newhouse, J. P. (szerk.): Handbook of health economics*. Elsevier. Amsterdam.
- KRONICK, R. ET AL. [1996]: Diagnostic risk adjustment for Medicaid: the disability payment system. *Health Care Financing Review*. 17. évf. 3. sz. 7–33. old.
- LAMERS, L. – VAN VLIET, R. [2004]: The pharmacy-based cost group model: Validating and adjusting the classification of medications for chronic conditions to the Dutch situation. *Health Policy*. 68. évf. 1. sz. 113–121. old.

- LAMERS, L. M. – VLIET, R. C. [2003]: Health-based risk adjustment Improving the pharmacy-based cost group model to reduce gaming possibilities. *European Journal of Health Economics*. 4. évf. 2. sz. 107–114. old.
- Magyar Köztársaság Kormánya. [2007]: T/4221. számú törvényjavaslat az egészségbiztosítási pénztárakról és a kötelező egészségbiztosítás természetbeni ellátásai igénybevételének rendjéről. T/4221.
- MANNING, W. – BASU, A. – MULLAHAY J. [2003]: *Generalized modeling approaches to risk adjustment of skewed outcomes data*. 293. 1–42. Cambridge, National Bureau of Economic Research. Technical Working Paper. Munkaanyag.
- MANNING, W. – BASU, A. – MULLAHAY, J. [2005]: Generalized modeling approaches to risk adjustment of skewed outcomes data. *Journal of Health Economics*. 24. évf. 3. sz. 465–488. old.
- MILGROM, P. – ROBERTS, J. [1990]: The efficiency of equity in organizational decision processes. *American Economic Review*. 80. évf. 2. sz. 154–159. old.
- MULLAHAY, J. [1998]: Much ado about two reconsidering retransformation and the two-part model in health econometrics. *Journal of Health Economics*. 17. évf. 3. sz. 247–281. old.
- NAGY B. [2006]: *A fejkvóta alapú forrásallokációs modell továbbfejlesztése - Elméleti alapok és gyakorlati lehetőségek a fejkvóta alapú forrásallokációs modell továbbfejlesztésére Magyarországon*. Egészségügyi és Stratégiai Kutatóintézet. Budapest.
http://www.eski.hu/new3/politika/zip_doc_2006/elmeleti_alapok.pdf
- NAGY B. – DÓZSA CS. [2002]: Az irányított betegellátási modellkísérlet tapasztalatai. *Egészségügyi Menedzsment*. 4. évf. 4. sz. 55–61. old.
- NAGY B. – DÓZSA CS. – BONCZ I. [2004]: A fejkvótaszámítás továbbfejlesztésének lehetőségei az irányított betegellátási rendszerben. *Egészségügyi Gazdasági Szemle*, 42. évf. 2. sz. 15–24. old.
- NAGY B. – SÍPOS J. – NAGY J. [2007]: Illusztrációk a fejkvótás forrásallokáció számításához Magyarországon – még csak a logikát ismerjük. *Informatika és Menedzsment az Egészségügyben*. 6. évf. 10. sz. 6–13. old.
- NEWHOUSE, J. P. [1996]: Reimbursing health plans and health providers: Efficiency in production versus selection. *Journal of Economic Literature*. 34. évf. 3. sz. 1236–1263. old.
- NEWHOUSE, J. P. [1998]: Risk adjustment: where are we now? *Inquiry*. 35. évf. 2. sz. 122–131. old.
- NEWHOUSE, J. P. ET AL. [1989]: Adjusting capitation rates using objective health measures and prior utilization. *Health Care Financing Review*. 10. évf. 3. sz. 41–54. old.
- RICE, N. – SMITH, P. [2001]: Capitation and risk adjustment in health care financing: an international progress report. *The Milbank Quarterly*. 79. évf. 1. sz. 81–113. old.
- RICE, N. – SMITH, P. [2002]: Strategic resource allocation and funding decisions. In: *Mossialos, A. et al. (szerk.): Options for Europe*. Open University Press. Buckingham.
- SALES, A. E. ET AL. [2003]: Predicting costs of care using a pharmacy-based measure risk adjustment in a veteran population. *Medical Care*. 41. évf. 6. sz. 753–760. old.
- SHEN, Y. – ELLIS, R. [2002]: How profitable is risk selection: a comparison of four risk adjustment models. *Health Economics*. 11. évf. 2. sz. 165–174. old.
- SMITH, P. – RICE, N. – CARR-HILL, R. [2001]: Capitation funding in the public sector. *Journal of the Royal Statistical Society*. 164. évf. 2. sz. 217–257. old.
- SZASZKÓ D. [2006]: Az OEP krónikus fekvőbeteg szakellátási kassza igénybevételének területi egyenlőtlenségei. *Magyar Epidemiológia*. 3. évf. 4. sz. 223–233. old.

- TAKÁCS E. ET AL. [2006]: Az OEP járóbeteg szakellátási kassza igénybevételének területi egyenlőtlenségei. *Informatika és Menedzsment az Egészségügyben*. 5. évf. 7. sz. 15–22. old.
- VAN DE VEN, W. – ELLIS, R. [2000]: Risk adjustment in competitive health plan markets. In: *Newhouse, J. P. – Culyer, A. J. (szerk.): Handbook of health economics*. Elsevier Science. London.
- VAN VLIET, R. [1992]: Predictability of individual health care expenditure. *The Journal of Risk and Insurance*. 59. évf. 3. sz. 443–460. old.
- WEINER, J. P. ET AL. [1996]: Risk-adjusted Medicare capitation rates using ambulatory and inpatient diagnoses. *Health Care Financing Review*. 17. évf. 3. sz. 77–99. old.
- WELCH, W. P. [2002]: Outpatient encounter data for risk adjustment: strategic issues for Medicare and Medicaid. *Journal of Ambulatory Care Management*. 25. évf. 3. sz. 1–15. old.
- ZHAO, Y. ET AL. [2005]: Predicting pharmacy costs and other medical costs using diagnoses and drug claims. *Medical Care*. 43. évf. 1. sz. 34–43. old.

Summary

This study aims to discuss the implementation of risk adjusted capitation in Hungary through a methodological overview and a practical example. The first part details the methodological questions of setting capitation by examining the formula development process; in the second part a selected patient group is analysed to demonstrate the possibilities of implementation. Both the methodological overview and the practical example justify the importance of statistical-mathematical methods, which have to be considered together with the experience of other disciplines during the budget setting process.

Esettanulmány egy nyugdíjas házaspár anyagi helyzetének alakulásáról 1990-ről 2007-ig

Dr. Csepely-Knorr András,

ny. c. egyetemi docens

E-mail: csepely@t-online.hu

A szerző egy már 1990-ben nyugdíjban lévő házaspár példáján mutatja be a nyugdíjasok anyagi helyzetének alakulását a rendszerváltástól 2007-ig. A nyugellátások reálértéke az első hat évben – egy választási évtől eltekintve – folyamatosan csökkent. Ezt tíz éven át növekedés követte, de ennek ellenére még az utolsó két évben, 2006/2007-ben is – differenciáltan – csak az 1990. évi reálértéket érte el. A növekedés a férj és a feleség esetében különböző volt. A nyugdíjak felhasználásánál döntő változások történtek a kiadások területén. A létfenntartási kiadásoknak tekinthető tételek közül három (fűtés, lakás, egészségügy) olyan mértékben emelkedett, hogy ezek 1990-ben még csak 18 százalékot kitevő aránya 2007-ig 44 százalékra nőtt, döntően lecsökkentve a nyugdíjból szabadon felhasználható hányadot. Statisztikai adatok bizonyítják, hogy a fenti tendenciák a több mint 1,6 millió főnyi öregségi nyugdíjban részesülőre egyaránt vonatkoznak. Bár a jelenlegi nyugdíjtörvény alapján a jövőben a nyugdíjak reálértékének szerény növekedésére lehet számítani, az élelmiszerek és az energiahordozók árának várhatóan továbbra is az átlagot meghaladó emelkedése a mindenkori nyugdíj összegéből csökkenő hányadot enged a nyugdíjasok számára szabadon választott kiadásokra (beszerzés, kultúra, közlekedés stb.) felhasználni.

TÁRGYSZÓ:

Nyugdíj.

Életszínvonal.

Az elemzésben szereplő házaspárból a feleség – egyetemi tanszéki adminisztrációjának vezetője – két évi kitelepítés, majd három gyermek felnevelése után 21 év munkaviszonnal a háta mögött 1986-ban, 55 évesen ment nyugdíjba. A férj – kutatóintézeti, majd vállalati főosztályvezető – 1989 végén 63. évében kérte nyugdíjazását, az egyetemi éveket is beszámítva 43 év munkaviszonnal. A nyugdíjazás előtti utolsó évben a fizetésből, prémiumból és jutalomból származó keresetük – egy főre számítva – közel kétszerese volt az országos havi átlagkeresetnek. Így a házaspár anyagi helyzete az aktív népesség átlagánál sokkal kedvezőbb volt, bár vagyon felhalmozását, de még jelentősebb tartalékképzést sem tett lehetővé. A saját feljegyzéseken alapuló longitudinális elemzésben a szerző nem elégszik meg a saját jövedelmi helyzet hosszú távú idősorán alapuló egyszerű bemutatásával, hanem ezt makrostatisztikai forrásokból származó adatokkal is igyekszik összehasonlítani.

1. A jövedelmek alakulása

A vizsgált első nyugdíjas év, 1990 folyamán a saját jogon nyugdíjazott munkásalkalmazotti állományúak átlagos havi nyugdíja 7 070 forint volt, amihez képest a feleség nyugdíjának összege ennek 84 százaléka, a férj nyugdíja pedig 196 százaléka volt, együttesen az egy főre jutó nyugdíjuk összege 43 százalékkal haladta meg a csoportjuk átlagát. Más közelítésben a házaspár havi nyugdíjának átlagos összege a KSH 1991-ben végzett felmérésében szereplő budapesti inaktív háztartások kiadásaihoz viszonyítva 43 százalékkal volt nagyobb. A nemzetgazdaságban az öt főnél több főt foglalkoztató vállalkozásoknál és intézményeknél alkalmazásban álló szellemi foglalkoztatottak havi átlagos nettó keresete 12 043 forint összegének a feleség nyugdíja 49,3 százalék, a férjé 115,2 százalék, együttes átlaga 82,3 százalék volt, vagyis, az előbbiekkal egybehangzóan, jelzi az átlaghoz képest lényegesen magasabb színvonalat, a házaspárnak a nyugdíjasok társadalmi rétegének legfelső 10 százalékának alsó részéhez tartozását, mégis lehetőséget ad a nyugdíjasokkal kapcsolatos néhány általánosításra is.

1990-ben, a házaspár nyugellátásának együttes összege 19 817 forint volt, aminek 70 százalékát a férj, 30 százalékát a feleség nyugdíja tette ki. Ebben az évben a férj szakértői tevékenysége révén egyszeri megbízásokkal havi átlagban 12 472 forint kiegészítő munkajövedelemhez jutott, ami 62,9 százalékkal, tehát közel kétharmadnyi-

val növelte jövedelmüket, és így az átlagos nyugdíjasoknál jóval nagyobb, közel kétszeres jövedelemhez juttatta a házaspárt.

A nyugellátás összege évenként a mindenkor előírt emeléssel növekedett. A feleség nyugdíjának alacsony összegét a férjéhez képest nagyobb arányban emelte a kormányzat 1991-ben, 1999-ben, majd 2007-ben, aminek eredményeként a korábbi arányok szerény mértékben (67-33 százalékra) változtak.

A házaspár mindkét tagját érintette 2002-ben egy egyszeri, az öregkori nyugdíjasok minden tagjára kiterjedő, 19 ezer forint összegű emelés. Ez gazdasági szempontból értelmezhetetlen volt, mivel egyrészt nem kapcsolódott a meglévő nyugellátási összegekhez, hanem egységesen részesült azonos összegben minden nyugdíjas, másrészt pedig nem épült be a nyugdíjakba, a következőkben nem képezte a mindenkori emelés részét. Egyértelműbb a 2003-tól fokozatosan belépő 13. havi nyugdíj pozitív hatása, mely értelemszerűen minden nyugdíjast a saját ellátásának megfelelő egyhavi összegben részesít a teljes felfutást jelentő 2006. év óta, azonban ez sem épül be az ellátás bázisába, és így nem képezi a következő évi nyugdíjemelés alapját, ami bizonytalanná teszi hosszú távú kiszámíthatóságát. Létjogosultságát ugyanakkor éppen a törvényt betervező és elfogadtató kormányzathoz közel álló szakemberek vitatják több-kevesebb gyakorisággal.

A feleség 1986-ban megállapított nyugdíjának korrekcióját a 2007-ben kapott és a jövőben is beépülő egyszeri 5 százalékos emelés képezte. Rendszeres jövedelemként jelent meg 1993-tól a feleség kitelepítéséért kapott kárpótlási jegyek életjáradékká válásából származó összeg, mely kezdetben a feleség nyugdíjának 15 százaléka volt. A későbbiekben ez az arány folyamatosan kissé csökkent (2006-ban már csak 11,7 százalék), mivel a mindenkori nyugdíjemelésnek az év eleji hányadát számítják be az életjáradék emelésébe, a keresetek vagy/és az infláció mértékének a tervezettnél nagyobb változásából eredő őszi korrekciók ezen a területen már elmaradnak.

A férj munkajövedelmének összege nominál értéken, öt éven át lényegében változatlan volt, reálértéken fokozatosan csökkent, majd 1995-től – egy-egy szakcikk tiszteletdíjától eltekintve – teljesen megszűnt, így ettől kezdve a házaspár jövedelme lényegében a nyugellátásra korlátozódott. A házaspár jövedelmi adatait a vizsgált kezdő- és záróéven kívül négy évenkénti átlagokkal az 1. táblázat foglalja össze.

A jövedelmek nominális értékének visszamenőleges adatai, a fogyasztói áraknak az évek során történt jelentős emelkedése miatt, nehezen értelmezhetők a mai árszínvonalon, ezért bemutatjuk a nyugellátások összegének alakulását az 1990. évi bázishoz viszonyítva a mindenkori fogyasztói árindex segítségével számított korrigált értéken (lásd a 2. táblázatot). Itt nem az átlagos inflációs indexet alkalmaztuk, mivel a nemzetgazdaság és a társadalom egészére vonatkozó átlagos fogyasztói árindex nem érvényes a nyugdíjasok csoportjára. Mások a mennyiségi és minőségi igények, valamint a lehetőségek, de mások a fogyasztás különböző kategóriáival kapcsolatos szükségletek is. Elég csak arra utalni, hogy különösen az utóbbi években gyors

ütemben növekvő árszínvonalú háztartási energia, fűtés termékcsoport iránti igény lényegében azonos az aktív keresők igényével, míg az átlagos áremelkedési ütemtől lényegesen elmaradó, sőt a legutóbbi években éppen csökkenő árú tartós fogyasztási cikkekre kisebb az igényük, de a beszerzésükre a lehetőségük is a nyugdíjasoknak. A számítások során ezért a Központi Statisztikai Hivatalnak mindezeket figyelembe vevő, az inaktív, illetve a nyugdíjas háztartásokra kiszámított fogyasztói árindexével dolgoztunk.

1. táblázat

*A házaspár nettó jövedelmének alakulása
(forint/hónap)*

Év	A férj	A feleség	Nyugellátás összesen	Feleség életjáradéka	Nettó munka jövedelem	Összes nettó jövedelem
	nyugdíja					
1990	13 878	5 939	19 817	–	12 472	32 289
1991–1994	21 694	9 868	31 562	1 007	8 896	41 465
1995–1998	41 490	21 795	63 285	2 850	878	67 013
1999–2002	72 465	34 721	107 186	4 931	1 066	113 183
2003–2006	113 767	53 744	167 511	6 987	–	174 498
2007	136 399	67 157	203 556	7 836	–	211 392

2. táblázat

A házaspár együttes havi nyugellátása és az infláció

Év	A házaspár nyugellátásának		Nyugdíjas fogyasztói árindex	Az 1990. évi nyugellátás reálérték meg- őrzéséhez szükséges ösz- szeg (forint)	A házaspár	A férj	A feleség
	tényleges havi összege (forint)	növekedése					
	Előző év=100		nyugdíjának tényleges összege a reálérték megőrzéséhez szükséges összeg százalékban				
1990	19 817	19 817	100,0	100,0	100,0
1994	40 945	106,6	146,5	48 844	83,8	82,3	87,4
1998	77 303	88,8	119,8	107 352	72,0	70,7	75,0
2002	132 685	71,6	40,0	150 318	88,3	84,8	96,3
2006	188 383	42,0	22,4	184 055	102,4	99,4	109,3
2007	203 566	8,1	10,7	203 749	99,9	95,6	110,0

Megjegyzés. A táblázatban felsorolt években mennyinek kellett volna lenni e a nyugellátás összegének ahhoz, hogy az 1990. évi nyugellátás reálértékét (a nyugdíjas indexszel korrigált értékét) megőrizze.

A teljes időszak első felében a fogyasztói árszínvonal – mindkét ciklusban – meghaladta a nyugellátás emelkedését, így az 1990. évhez képest annak reálértéke csökkent és 1998-ban már csak 72 százalék volt. Ezt követően, egyebek mellett, az 1997-ben konszenzussal elfogadott nyugdíjtörvény eredményeként, megfordult a helyzet és a nyugdíjak emelésének üteme lényegesen meghaladta a mindenkori áremelkedést, aminek következtében a reálérték dinamikusan nőtt és 2006-ban már ismét elérte – szerény mértékben meghaladta – az 1990. évi szintet, amit 2007. évben már ismét visszaesés követett. Ettől eltérő, a nyugdíjasok helyzetét kedvezőbben bemutató eredményre jutott az a jelentés, amely 2006-ban „Biztonságosabb idősor, felzárkózó nyugdíjak” címmel készült a kormány részére, mivel a számítások során nem a nyugdíjasok társadalmi csoportjánál érvényesülő, hanem az ettől rendre elmaradó átlagos árszínvonal-emelkedéssel számolt, azonban 1989-hez viszonyítva így is csak 11,5 százalék reálérték-emelkedést prognosztizált 2006-ig. A különbséget a számítások során az árszínvonal-emelkedés területén alkalmazott két eltérő indexsor indokolja, melyek között 2007-ig kerekén 10 százalékpont különbség alakult ki. Figyelemre méltó, hogy a feleségnek a mindenkori átlagos női nyugdíj körül mozgó ellátása az időközben történt külön emelések eredményeként valóban meghaladta reálértékben az 1990. évi szintet, míg a férj magas nyugdíja, ha csekély mértékben is, de még akkor is elmaradt attól, sőt 2007-ben ismét visszaesett az 1990. évi reálérték 95,6 százalékára..

A nyugdíjak reálértékének változása a ciklusokon belül igen különböző volt, amit a 3. táblázat adatai mutatnak. Az első két ciklusban történt jelentős visszaesésen, majd a második kettőben tapasztalt növekedés során egyaránt volt az általános tendenciától eltérő változás.

3. táblázat

A nyugdíjak reálértékének ciklusokon belüli változása

Ciklus	1991–1994	1995–1998	1999–2002	2003–2006	2007. évben
	évek közötti ciklus (százalékpont)				
1.	-8,6	-9,4	+2,5	+4,8	-2,5
2.	-6,9	-6,6	+0,4	+1,8	
3.	-3,8	+0,1	+3,6	+5,1	
4.	+3,1	+4,1	+9,8	+2,4	
<i>Ciklusban összesen</i>	<i>-16,2</i>	<i>-11,8</i>	<i>+16,3</i>	<i>+14,1</i>	

Az első két ciklusban a reálérték jelentős visszaesése fokozatosan enyhülő mértékben ment végbe és a ciklus utolsó évében – a választások előtt érthető módon – életszínvonal-javító intézkedések között a nyugdíjak értékromlását is megállították,

sőt szerény javításra is sor került. A harmadik és a negyedik ciklus végig pozitív előjelű éveiben sem volt egyenletes a fejlődés. A harmadik ciklusban a második év hozott megbicsaklást, különben az első és a harmadik évben egyaránt 3 százalékpont körül volt a javulás, amit a negyedik, választási évben – az új kabinet egységes 19 ezer forint összegű juttatása figyelmen kívül hagyása nélkül is – megkétszerezett a leköszönő kormányzat. A negyedik ciklus a korábbiaktól teljesen eltérő nyugdíjpolitikát érvényesített. A már említett, még a választás évében nyújtott egyszeri 19 ezer forintot követően az első évben – kapcsolódva a rendkívüli bérfeljesztésekhez – az előző ciklus utolsó évéhez hasonló nagyságú nyugdíjemelést biztosítva közel 5 százalékponttal emelte a nyugdíjak reálértékét. Ezt a nyugdíjasok nézőpontjából egy szerény eredményt hozó év követett, majd a harmadik évben ismét jelentős mértékben 5,1 százalékponttal nőtt a reálérték. Végül a negyedik évben – minden eddigitől eltérően – a választási év szerény, 2,4 százalékpontnyi reálérték-javulást hozott. A bázisévet követő első két ciklus 30 százalékpontot meghaladó reálérték-csökkenését követően a második két ciklusban hasonló mértékű, 16,3, illetve 14,1 százalékpontnyi növekedéssel végül is lényegében a bázisév reálértékének színvonalát élvezhették a már 1990-ben nyugdíjban levők. Hozzá kell tenni ehhez azonban azt, hogy ez az eredmény csak a 13. havi nyugdíj bevezetésének volt köszönhető. A számítások azt mutatják, hogy a 13. havi nyugdíj nélkül a negyedik ciklusban a nyugdíjak reálértéke mindössze 7,6 százalékponttal nőtt volna és így 2006-ban az 1990. évhez képest még mintegy 5 százalékponttal elmaradt volna a reálérték.

A mindenkori nyugdíjösszeg nagyságának megítélésének egyik aspektusa az eltelte időszak alatt a reálérték előbbiekből bemutatott változása, másik lehetséges közelítés az aktív népesség keresetével történő egybevetés. Erre a nemzetgazdaságban az öt főnél nagyobb létszámot foglalkoztató vállalkozásoknál és intézményeknél alkalmazásban álló aktív keresők nettó keresetének a 4. táblázat szerinti adatai adnak lehetőséget.

Az aktív keresők és a nyugdíjasok ellátásának egybevetése előtt meg kell jegyezni, hogy a foglalkoztatottak havi átlagos nettó keresetének reálértéke 1990 és 2006 között – a fenti adatokból számítva – átlagosan 26,8 százalékponttal, míg ezen belül a szellemi munkát végzőké 30 százalékpontot meghaladó mértékben nőtt. Ennek következtében a házaspár nyugellátása az aktív csoport nettó keresetéhez viszonyítva az évek során folyamatosan csökkent, kezdetben gyorsabb, majd lassúbb ütemben. Az első négy évben a nyugdíjmelés mindvégig jelentősen elmaradt a keresetek, és főként a korábbi aránytalanságokat csökkenteni kívánó bérpolitika hatásaként, a szellemi foglalkoztatottak legalább viszonylag gyorsabban növekvő nettó keresete mögött, és így az átlaghoz viszonyítva 9 százalékpont, a szellemi foglalkoztatottakhoz viszonyítva pedig 13 százalékpontnyi veszteséget szenvedett. Ezen a második és a harmadik kormányzati ciklus is, csak annyiban változtatott, hogy a további visszailleség, most már a nyugdíjtörvényben foglalt indexálásnak megfelelően, az aktív kere-

sökhöz képest, az infláció és a keresetek növekedésének 50-50 százalékos figyelembe vételének következtében, kisebb mértékű volt a korábbiaknál. Figyelemre méltó, hogy a 2002 és 2006 közötti években a 13. havi nyugdíj fokozatos belépése ellenére tovább folytatódott, a házaspár nyugdíjának csökkenése igaz csak 0,6 százalékponttal az összes keresőhöz képest, míg a szellemi foglalkoztatottakhoz viszonyítva a korábbi 65,3 százalék szinten tudott maradni. A szellemi foglalkoztatottak reálkeresetének 2007. évi igen kedvezőtlen alakulása, a nominális érték 1,8 százalékos növekedése az összes foglalkoztatott nettó keresetének 2,9 százalékos növekedésével szemben (8,0 százalék átlagos infláció mellett), valamelyest javított ugyan a nyugdíjasok viszonylagos helyzetén, ami akár sikerként is elkönnyvelhető lenne, ha a nyugdíjak 8,1 százalékos növekedése mellett a nyugdíjas fogyasztói árindex nem lett volna 10,7 százalék, így a vásárlóerő nem csökkent volna.

4. táblázat

A havi nettó keresetek és a nyugdíjak változása

Év	Az összesen	A szellemi foglalkoztatottak	A férj	A feleség	A férj	A feleség	Együtt
	foglalkoztatott havi átlagos nettó keresete		egy főre jutó ellátása		a nyugellátás az aktív szellemi foglalkoztatottak nettó keresetének hányadában		
	(forint/hónap)				(százalék)		
1990	10 108	12 043*	13 878	5 939	115,2	49,3	82,3
1994	22 992	29 576	28 153	12 792	95,2	43,3	69,2
1998	44 383	58 072	53 165	24 138	91,6	41,6	66,6
2002	77 607	101 658	89 316	43 369	87,9	42,7	65,3
2006	110 896	144 244	128 102	60 281	88,8	41,8	65,3
2007	114 112**	146 940**	136 399	69 157	92,8	47,1	69,9

* Becsült adat.

** KSH előzetes adatai 2008.02.20.

Forrás: KSH-STADAT A foglalkoztatottak átlagos nettó keresete a nemzetgazdaságban.

Az eddigiekben rendre magasnak ítélt nyugellátással kapcsolatban érdemes egy tanulmány megállapításaival szembeállítani. Az „Időskorúak Magyarországon” című tanulmány (KSH [2004]) sok szempontból részletesen vizsgálja a különböző csoportok helyzetét és többek között bemutatja a vizsgálatba vontak szerint a megélhetéshez szükségesnek tartott havi összeget. Ezek szerint a Budapesten élő idősök véleménye azt mutatja, hogy a tényleges 60 ezer forint/hó egy főre jutó jövedelemmel szemben „átlagos szintű” megélhetéshez 2002-ben 71 ezer forint, „jó színvonalhoz” 107 ezer forintra lett volna szükség. A házaspárnak a tényleges átlaghoz viszonyítva

igen jónak ítélt egy főre jutó nyugdíja, a tanulmányban foglaltak szerint az átlagos színvonalhoz sem volt elegendő és az előbbieken bemutatottak szerint azóta sem változott lényegesen a helyzet.

A fogyasztási javak és szolgáltatások arányának és egyúttal árszínvonalának a nyugat-európai szinthez közelítésével kapcsolatos prognózisokhoz és valós tendenciákhoz kapcsolódva a 5. táblázatban a házaspár együttes jövedelmét, a Magyar Nemzeti Bank által jegyzett deviza-középfolyam éves átlagával számolva, ecu/euro értéken mutatjuk be. Ezt elsősorban nem a jövedelmek nagyságának bemutatása érdekében tesszük, hanem az időszak során bekövetkezett változás tendenciájára kívánjuk felhívni a figyelmet.

5. táblázat

A házaspár havi nettó jövedelme ecu/euróban

Jövedelem típusa	1990. év	1991–1994	1995–1998	1999–2002	2003–2006	2007. év
	évek átlaga					
Férj nyugdíja	172	203	205	287	447	543
Feleség nyugdíja	74	92	108	138	211	267
Nyugellátás együtt	246	295	313	425	658	810
Feleség életjáradéka	–	9	14	20	28	31
Nettó munkajövedelem	154	84	4	4	–	–
<i>Összes nettó jövedelem</i>	<i>400</i>	<i>388</i>	<i>331</i>	<i>449</i>	<i>686</i>	<i>841</i>

Csak a nyugellátás összegére korlátozva a vizsgálatot, az első két ciklusban stagnálást, majd a második két ciklusban gyorsuló növekedést állapíthatunk meg. A változáshoz két megjegyzést kell fűzni. Az első az ellátás nagyságának értékelése nemzetközi összehasonlításban, a második pedig a nemzeti valuta és az euró értékárnyának változása.

Az átlagos nyugdíjak alakulásáról az Európai Unió tagországaiban nem áll rendelkezésre összehasonlítható idősor. Ennek hiányában a mindenkori minimálbér alakulásához viszonyítjuk a házaspár nyugdíját. A tizenöt tagországból kilenc országban törvényben volt meghatározott minimálbér. Ennek nagysága a kilenc ország átlagában 2000-ben 852 euró volt havonként, majd évenként eltérő mértékű növekedés után ért el 2005-re 1 025 euróra. A hazai viszonyok között tehát a házaspár magasan tekinthető egy főre jutó nyugdíja euróban számolva, dinamikus növekedése ellenére, csak egynegyede, egyharmada az átlagos uniós minimálbérnek. Annak ellenére van így, hogy éppen ebben az időszakban erősödött a nemzeti valuta jelentős – sok szakértő szerint irreális – mértékben az euróhoz képest.

Ez utóbbi megítéléséhez ismét egy összehasonlítással élünk uniós adatok segítségével (ECU/EUR exchange rates versus national currencies – Eurostat). Az 1998 és 2006 közötti években Magyarországon összesen 67,2 százalék infláció mellett a forint csak 9,7 százalékkal romlott az euróhoz képest. Mi történt ugyanakkor a többi csatlakozó országban? Példaként: noha Csehországban a fogyasztói áremelkedés csak 19,2 százalék volt ebben az időszakban, de ugyanakkor az euró értéke 36,40 koronáról 28,34 koronára, tehát 22,1 százalékkal csökkent, Lengyelországban 36,6 százalék infláció mellett (3,92-3,90) lényegében változatlan maradt a zloty és az euró árfolyama, Szlovákiában 71,7 százalék fogyasztói áremelkedés mellett az euró árfolyama 39,63 koronáról 37,23 koronára mérséklődött, míg Szlovéniában az infláció 55,3 százalék volt és az euró árfolyama 186,37 tolarrról 29,3 százalékkal 240,89 tolarra nőtt. Kisebb vagy nagyobb mértékben ugyanez állapítható meg a többi csatlakozó országról is, azaz a nemzeti valuta inkább nagyobb, mint kisebb mértékben felértékelődött az euróhoz képest. Az említett szakmai kérdőjelek ellenére sem tételezhető fel, hogy minden ország központi bankja és pénzügyi kormányzata szakszerűtlen vagy felelőtlen árfolyam-politikát folytatott volna 1998 és 2006 között, inkább a következő megállapítás lehet helyénvaló.

A gazdasági növekedéssel kapcsolatban végzett vizsgálataik során tíz évvel ezelőtt állapította meg *Artner Annamária* és *Inotai András* kutatása, hogy minden országban két részre kell választani a GDP növekedésének vizsgálatát: a hagyományos mennyiségi növekedést tükrözi az „igazolható” reálérték változása, míg a minőségi változást, a termelékenység javulását, a ráfordítások hatékonyságának változását bizonyítja a nemzeti valuta felértékelődéséből származó GDP-növekedés. Mint a kutatók megállapítják: „... adott ország GDP/főben mért és dollárban kifejezett felzárkózása csak kis részben függ a felzárkózó és a mércének tekintett ország (régión) között az előbbi javára megnyilvánuló növekedésiütem-különbségtől. Meghatározó szerepet az árfolyamváltozás, vagyis a felzárkózó ország nemzeti valutájának felértékelődése játszik... Joggal vélelmezhető, hogy a közép-európai tagjelöltek felzárkózásában is meghatározó szerepet fog játszani az összehasonlítható valutában (dollár, márka, ecu) kifejezhető felértékelődés, ami a sikeres modernizációt az esetek túlnyomó többségében kísérni szokta.” (*Artner–Inotai* [1997]) A kutatók előrejelzését messzemenően igazolták a következő tíz év tényleges adatai a régió minden országában.

Összegezve tehát a házaspár közel két évtized során elért jövedelmének alakulásából több lényeges következtetésre és megállapításra juthatunk, melyek következményeivel minden nyugdíjasnak számolnia kell:

- a nyugdíjazás utáni első években még elég gyakori a nyugdíjat kiegészítő munkajövedelem, ez azonban fokozatosan elmarad, így a háztartás-gazdálkodás bevételi oldala leginkább a nyugellátásra korlátozódik (így volt ez az elemzésben bemutatott házaspár esetében is, a

kiegészítő munkajövedelem nyugdíjat elérő kétharmados mértéke azonban természetesen nem általános.);

– a nyugdíjak reálértéke az első két ciklusban 1990-hez viszonyítva összesen 28 százalékot veszített, majd a második két ciklusban érte el ismét az 1990. évi reálértékét;

– egy tanulmány szerint (*Artner–Inotai* [1997]) a GDP növekedése a vizsgált időszak alatt 32,6 százalék volt, amivel szemben, a nyugdíjak reálértéke még az átlagos inflációt figyelembe véve is csak 11,5 százalékkal nőtt, valójában éppen csak a bázisév szintjén volt az időszak végén, míg a nemzetgazdaságban foglalkoztatottak átlagos nettó keresetének reálértéke ugyanakkor 26,8 százalékkal, ezen belül a szellemi foglalkoztatottaké 30 százalékpontot meghaladó mértékben nőtt;

– az aktív dolgozók nettó keresetéhez viszonyítva a házaspár nyugellátásának tényleges összege 98 százalékról 85 százalékra, míg a szellemi foglalkozásúak átlagához viszonyítva 82 százalékról 65 százalékra esett vissza;

– mindezek ellenére az Európai Unió valutájára átszámítva a házaspár egy főre jutó havi átlagos nyugdíja 123 euróról, elsősorban a második nyolc év során több mint megkétszereződött, 329 euróra majd 2007-ben 405 euróra nőtt, ami kizárólag a forintnak az euróhoz viszonyított felértékelődéséből származott.

2. A kiadások változása

Mint láttuk a házaspár bevételei között, nyugdíjas éveik elején jelentős, bár fokozatosan csökkenő szerepet játszottak a kiegészítő munkajövedelmek. Az első négy év után azonban ezek gyakorlatilag megszűntek és a kiadások fedezetét elsősorban a nyugdíjak szolgálták. A jövedelmek ilyen változása természetesen sok tekintetben befolyásolta a házaspár gazdálkodását, kiadásaik nagyságát és összetételét. Ennek megfelelően az idősor vizsgálata során a bázis- és az utolsó év mellett külön kiemeltük az 1995. évet, mint a kizárólag nyugdíjjövedelemből álló első esztendő. A kiadásokat a családi háztartási napló bontása alapján 11 csoportra osztottuk. Ezek az 6. táblázatban bemutatott csoportok tartalmukban eltérnek a statisztikában a lakosság fogyasztásának bemutatásához általában alkalmazott részletezéstől, de sok tekintetben így is alkalmasak azokkal való összehasonlításra.

A kiadásokon belül a két nagy csoport, az adott körülmények között (életkor, egészségi állapot, lakásviszonyok) szükséges létfenntartási költségek és a házaspár szabad elhatározása alapján választott vásárlások, kiadások arányaiban igen nagy

változás történt az idő során. A munkajövedelemmel rendelkező években a létfenntartási kiadások aránya ugyan – a bevételek fokozatos csökkenésével párhuzamosan – nőtt, de mindvégig az összes kiadás 50 százaléka alatt maradt. A kizárólag nyugdíjra alapozott bevételek kezdetétől már meghaladta az 50 százalékot és az arányokat tekintve viszonylag gyorsan emelkedett, az utolsó években elért közel 70 százalékig. A családi gazdálkodásban ez a szabad mozgástér beszűkülését, a kiadások terén a de-termináltság egyre nagyobb mértékét jelenti.

6. táblázat

A kiadások összetételének változása

Kiadási tételek, termékcsoportok	1990-ben	1991–1994	1995-ben	1995–1998	1999–2002	2003–2006	2007-ben
	évben átlagosan (százalék)						
1. Élelmiszer és háztartási vegyi áru	12,6	14,4	19,2	19,6	18,0	15,9	14,0
2. Fűtés, háztartási energia	2,8	4,1	6,6	8,0	9,4	10,1	11,2
3. Lakással kapcsolatos állandó költségek	14,0	12,7	12,2	13,6	16,0	21,4	18,9
4. Egészségügyi kiadások	1,4	3,5	10,5	7,5	7,9	12,4	13,9
5. Lakás-karbantartás és kisebb szolgáltatások	9,9	11,6	7,1	7,3	9,4	7,8	4,4
<i>Létfenntartási kiadások együtt</i>	<i>40,7</i>	<i>46,3</i>	<i>55,6</i>	<i>56,0</i>	<i>60,7</i>	<i>67,6</i>	<i>66,4</i>
6. Élvezeti cikkek	9,9	6,0	4,5	5,6	6,9	4,8	4,1
7. Ruházati és tartós fogyasztási cikkek	14,6	11,9	7,8	6,2	6,6	4,2	8,1
8. Családdal kapcsolatos kiadások	13,1	11,8	9,2	11,1	9,0	11,0	8,7
9. Kultúra, reprezentáció	9,4	10,5	10,9	8,2	7,2	7,2	9,4
10. Közlekedés, gépkocsi fenntartása	11,1	12,4	10,6	11,8	8,1	3,5	6,1
11. Egyéb kiadások	1,2	1,1	1,4	1,1	1,5	1,7	1,2
<i>Szabadon választott kiadási tételek</i>	<i>59,3</i>	<i>53,7</i>	<i>44,4</i>	<i>44,0</i>	<i>39,3</i>	<i>32,4</i>	<i>33,6</i>
<i>Összes kiadás</i>	<i>100,0</i>	<i>100,0</i>	<i>100,0</i>	<i>100,0</i>	<i>100,0</i>	<i>100,0</i>	<i>100,0</i>

1. *Élelmiszerek, háztartási vegyiáruk fogyasztása.* Ebbe a csoportba tartoznak az étkezéshez felhasznált nyers és feldolgozott élelmiszerek, valamint a tisztálkodáshoz és a háztartáson belüli tisztításhoz (takarítás, mosogatás, mosás stb.) szükséges háztartási vegyi-árúk. Ennek a termékcsoportnak az árindexe általában kisebb mértékben emelkedett, mint az általános fogyasztói árindex, így változatlan mennyiségű és minőségű fogyasztás esetén a termékcsoportra fordított kiadások arányának az összes kiadásokon belül csökkennie kell.

A házaspár esetében ez a tendencia csak 1995 után következett be, bár korábban a munkajövedelemmel növelt magas bevételek szabadabb költést tettek lehetővé, ez a többlet messze elmaradt a többletbevétel nagyságától, így alacsonyabb kiadási hányadot adott. A későbbiekben azonban már érvényesült az árváltozások különbségéből adódó tendencia. Az 1990-ben a termékcsoportra fordított 3 892 forint kiadást a termékcsoport árváltozásának mértékével indexálva 2006-ban 30 380 forint összeg lett volna az indokolt kiadás. A valóságban azonban ez csak 23 170 forint, tehát a változatlan mennyiségű és minőségű fogyasztásnak csak 76,3 százaléka. Az első kizárólag nyugdíjas évek, 1995-nek a fogyasztása 9 813 forint/hó volt, amihez képest a termékcsoport árindexével növelt indokolt fogyasztás 25 179 forint lett volna 2006-ban, de ez a valóságban ennek csak 92,1 százaléka, 23 170 forint/hó volt, vagyis változatlan áron mérve ebben az esetben is csökkent a termékcsoportra fordított kiadás.

A fogyasztás visszafogása azonban sem az első, sem a második esetben nem a korábban sem túlzott élelmiszermennyiség csökkentésében nyilvánult meg, hanem egyrészt a a minőségi áruk felhasználásának jelentős mérséklésében, másrészt pedig az olcsóbb értékesítési helyek, az azonos termékért fizetendő alacsonyabb árak felkutatásával. Az élelmiszerek és a háztartási vegyi áruk között egyaránt a korábban inkább megengedett külföldi termékek, vagy az évek folyamán megjelent új termékek is egyre kevésbé kerültek a vásárlói kosárba. A nyugdíjas szabadidejének felhasználásával – amíg azt egészsége engedi – megteheti, hogy nem néhány, a lakásához legközelebb fekvő üzletben vásárol, hanem tapasztalatait és a mindenkori hirdetések kínálatát is figyelembe véve igyekszik, a legolcsóbb terméket megtalálni és megvásárolni. Így próbálja a piacok és az egyre nagyobb számban megnyíló bevásárlóközpontok kínálatát kihasználni, már amennyire számára ezt megközelíthetőségük lehetővé teszi. Végül, de nem utolsósorban, számottevő megtakarítás érhető el a napi vásárlás helyett az egy-, kéthetenkénti vagy egy-egy termék esetében akár havonkénti vásárlás előnyeinek kihasználásával. A háztartási vegyi árúknál pedig főként a kipróbálás során megfelelőnek minősített közepkategóriás, többnyire hazai termékek használata révén tudott a házaspár megtakarítást elérni. Mindezek együttesen eredményezték, hogy az első ciklusban a házaspár kiadásainak még csak 14,4 százalékát fordította a termékcsoportra, ami a második ciklusban már 19,6 százalék volt, majd a harmadikban 18,0 százalékra, a negyedikben pedig 15,9 százalékra mérséklődött. Ez utóbbi azonban a továbbiakban aligha tartható, mivel egyrészt már az utolsó két évben (2006–2007) az átlagos inflációt meghaladó mértékű volt a termékcsoport árindexének növekedése, másrészt pedig – részben az Európai Unió előírásai miatt is – a termékek kulturáltabb csomagolása, tárolása, terítése a tartalmában változatlan termékek árának is további jelentős emelkedését vonja maga után, ami összességében azt jelenti, hogy az egyéb termékeknél részben már megvalósultakhoz hasonlóan az élelmiszereknél is az Unió országaiban tapasztalható árszínvonal és arányok kialakulásához fog vezetni, megemelve ez-

zel a családok költéséből a termékcsoportha fordított hányadot. Ennek a megállapításnak látszólag ellentmondanak a 2007. évi adatok, de ezeknek sajátos, egy évben érvényesülő okai és hatásai nem általánosíthatók, a termékcsoportha fordított kiadások arányának növekedése várható.

2. *Fűtés, háztartási energia.* A lakásban gázüzemű cirkófűtés működik, ugyancsak földgáz biztosítja a fürdőszobai és konyhai melegvíz-szolgáltatást és adja a főzéshez szükséges energiát. Az elektromos áramszolgáltatás a világítás mellett a háztartási gépek (hűtő és fagyasztó szekrény, mikrohullámú sütő, mosógép, vasaló és konyhai kisgépek), valamint a szórakoztató elektrotechnikai eszközök energiaigényét elégíti ki.

A fogyasztás mindkét energiahordozóból viszonylag stabil, a tudatos takarékoság ellenére – mivel korábban sem volt pazarlás – lényeges változás nem mutatható ki. A gázfogyasztás évi 1800 és 2000 köbméter között váltakozott az elmúlt 18 évben, elsősorban a téli hideg mértéke és a tél hossza befolyásolt a fogyasztás mennyiségét. A fűtés nélküli nyári hónapokban a havi gázfogyasztás 50-70 köbmétert ér el, majd fokozatos átmenet után, a leghidegebb téli hónapokban havi 250-300 köbméter a régi építésű, így magas belmagasságú 300 légköbméteres lakás teljes fűtésakor. A villamos áram évi 1700-1900 kilowattóra felhasználásában nincs ilyen jelentős különbség: a nyári hónapokban átlagosan 120-140, az őszi-téli hónapokban 150-170 kilowattóra a fogyasztás, ami azt is jelenti, hogy az áramfelhasználás nagyobb hányadát a gépek-eszközök fogyasztása teszi ki és csak kisebb hányadát a javarészt energiatakarékos izzókkal biztosított világítás fogyasztása.

A kiadások között a költségsoport súlya kezdetben szerény volt, az első ciklusban 4,1 százalék, majd a másodikban megkétszereződve 8,0 százalék, a harmadik és a negyedik ciklusban kisebb növekedéssel 9,4 százalék, illetve 10,1 százalék, majd az utolsó, a 2007. évben az arányoknak az előbbi termékcsoportha tapasztalt eltérése ellenére már 11,2 százalék volt. Az arányok a lényegében változatlan szintű felhasználás mellett nemcsak a jövedelmek csökkenése miatt növekedtek, hanem a termékcsoportha árszínvonalának az átlagot jóval meghaladó mértékű emelkedése miatt is. Jól érzékelteti ezt mindkét energiahordozónál egy-egy adat. Az 1990. év átlagának 4,03 forint/köbméter gázárával szemben 2007-ben ennek már kereken huszonötszörösét, az elektromos áram 1,82 forint/kilowattóra árával szemben annak több mint húszszorosát kellett a házaspárnak fizetnie (gáz-ártámogatásban a határértéket éppen meghaladó „magas” jövedelmük miatt nem részesülnek).

A háztartási energia termékcsoportha mutatja talán legtisztábban azokat az ellentmondásokat, amelyek az elmúlt évtizedek torz árrendszerében és ezt mintegy ellentételezve az irreálisan alacsony jövedelmekben mutatkoztak meg. Ezek fokozatos feloldása elsősorban a lakosságnak jelentett és jelent ma is rendkívül nagy, kötelező áldozatvállalást az árak európai szintre emelkedése és egyidejűleg a jövedelmek szerény mértékű növekedése, a régi nyugdíjasok, így a házaspár esetében még ettől is elmaradó, reálértékben lényegében stagnálása miatt.

3. *A lakással kapcsolatos állandó költségek.* A házaspár 79 m² alapterületű kétszobás lakásban él. A lakás 2000 közepéig önkormányzati tulajdonban volt, ekkor privatizálták, aminek keretében azt a házaspár leánya vette meg holtig tartó haszonélvezeti jogot biztosítva szüleinek. Így ebbe a költségcsoportba kezdetben a lakbér, a víz- és csatornadíj, 1997-től további tételként a szemétszállítás díja, került, majd ezek helyébe lépve 2000-tól, a privatizáció után a lakást terhelő közös költség követte. Ebbe a csoportba került még a telefon költsége, az AM-mikro antenna-csatlakozás, illetve az ezt felváltó, egyidejű internetkapcsolatot biztosító T-Home használat díja.

A tételek között a legnagyobb, az adott körülmények között a házaspár gazdálkodásától független, a lakás közös költségének alakulása, melyet egyrésztől a különböző szolgáltatók (víz, csatorna, elektromos áram, szemétszállítás) díjai, másrészt a lakóközösség évi közgyűlése határoz meg. Ez utóbbiba tartozik a ház takarításának költsége, a közös képviselő díja, valamint a ház karbantartásához, a szükséges részleges felújításhoz biztosított tartalék összege. Ez utóbbiból a privatizációt követő második évtől (amikorra már valamennyi tartalék képződött) a lakóközösség döntése alapján – az elmúlt évtizedekben elmaradtak fokozatos pótlásaként – minden évben megvalósult a tetőtértől kiindulva a feltétlenül szükséges karbantartáson túl egy-egy szerkezeti elem (tetőhéjazat, eresz- és lefolyócsatorna, gázvezeték) részleges vagy teljes cseréje, amiket a továbbiakban kell még az elektromos vezetékek, a külső és belső homlokzatok, nyílászárók felújításának, cseréjének követnie. Az így havonként fizetendő közös költség összege 2007-ben 263 forint/négyzetméter, ezen belül a felújítási alap része 85 forint/négyzetméter, ami lényegesen magasabb, mint az országos átlag, de ez teszi lehetővé a 21 lakásos épület belátható időn belüli megfelelő, felújított állapotba hozását. A KSH adatai szerint ugyanakkor a társasházi közös költség országos átlaga (50 négyzetméteres lakás díjából visszaszámolva) 218 forint/négyzetméter volt. A két adat közötti 20 százalékos eltérés nem jelent a házaspárnak különleges terhet, bár a teljes összeg már az együttes havi nyugdíj közel 15 százaléka, így arányát tekintve közel európai szint.

A telefon-előfizetési díj, arányát tekintve, csökkenő tendenciájú az igénybevett kedvezményes, de egyidejűleg az alacsony díj miatt korlátozottan használható szolgáltatás miatt. Ugyanakkor jelentősen nőtt a tévé üzemben tartásához szükséges csatlakozás díja, kétségtelenül szélesebb körű szolgáltatással. Új elemként lépett be az internetkapcsolat előfizetése, mely a korszerű élet egyre nélkülözhetetlenebb eleme még a nyugdíjas családoknál is, kiválthatja a nyomtatott sajtóra fordított kiadásokat és hozzátartozik az egyre inkább a lakásba visszaszorulás mellett a világot befogadni kívánó élethez.

A kiadások aránya így az 1991/1994 évi 12,7 százalékos átlagról fokozatosan, de gyorsuló ütemben emelkedett a harmadik ciklussal bezárólag 16,0 százalékra, majd az utolsó négy évben 21,4 százalékra.

4. *Egészségügyi kiadások.* Az arányaiban 1990-hez viszonyítva közel tízszeresére növekvő tétel az egészségügy átalakításának időszakában külön figyelmet érdemel. Ezek a kiadások a nyugdíjas korúaknál értelemszerűen magasabbak, mint az átlagos aktív korúaknál, de a csoporton belül is az idő előrehaladtával érthetően növekvő tendenciájú az öregedő szervezet egészségi állapotának fokozatos romlása okán. Egyik oldalról tehát objektív tendenciáról van szó, ami jobbra független a gazdasági környezet hatásaitól. Az idősödő ember egyre többször kénytelen igénybe venni orvosi szolgáltatásokat és ezekkel párhuzamosan nagyobb gyógyszer-mennyiséget, amik természetesen növekvő költségekkel járnak.

Az orvosi költségek között az egészségügyi pénztár által nem finanszírozott ellátások (fogorvosi ellátás több tétele), a paciens által választott szakorvos díjazása (hálapénz), a házi orvos és asszisztense karácsonyi megemlékezése, valamint 2007-ben már a vizitdíj került a csoportba. Az 1990. évi 0,7 százalékos arány az összes költségen belül a még kettős jövedelem mellett kialakult egy évi esetleges érték reálisabb, de még mindig igen alacsony: az 1991/1994. évek átlagának 2,0 százaléka. A kilencvenes évek második felének álagában az összes kiadáson belül 4,7 százalékos, a századforduló első éveiben tovább emelkedve 6,6 százalékos, majd az elmúlt utolsó négy évben már 8,5 százalékos és 2007-ben 8,9 százalékos volt a kimondottan orvosi ellátásokra fordított kiadás.

Ugyancsak dinamikus volt a költségek arányának növekedése a gyógyszereknél. A kilencvenes évek első négy évének 2,0 százalékos átlagáról négy évenként mintegy 2 százalékpont körül nőtt a költségek aránya és ért el az utolsó négy évben 8,5 százalékos. Kétségtelen, hogy korábbi gyógyszereket korszerűbbre cserélve a jobb hatást magasabb árral kellett megfizetni, de a házaspár esetében nem ez volt a növekedés fő oka, ezt mutatja az 1995-ig visszanyúló, tíz rendszeresen szedett, alkalmazott gyógyszer áralakulása. Ezek között az egyszerű fájdalomcsillapítótól a nyugtatón, altatón át az érrendszer karbantartását szolgáló és a bőrfelület belső okokból fakadó krónikus betegség tüneteit korlátozó, így támogatott és nem támogatott szerek egyaránt szerepelnek. Az így minden nap szedett gyógyszerekre fordított kiadás 1995-ben 785 forint, 2007-ben már ennek tizenkétszerese, 9 440 forint volt. A ciklusokat egyenként vizsgálva egyértelműen az állapítható meg, hogy azok mindegyikében messze meghaladta a gyógyszerek áremelkedése az éppen aktuális átlagos infláció mértékét. Így érthető, hogy „fogyasztásváltozás” nélkül is a kilencvenes évek első négy évének átlagához képest az összes kiadáson belül több mint négyszeresére emelkedett a gyógyszerekre fordított kiadások összegének aránya. Ez a „gyógyszercsomag” természetesen nem azonosítható a nyugdíjasok átlagos felhasználásával, bár nagyon valószínű, hogy több betegség ellen használt többféle gyógyszer összesített költsége más felhasználóknál sem mutatna alapvető eltérést az előbbiekhöz képest. Az egészség megtartására és védelmére fordított kiadások aránya a házaspár összes kiadásán belül többszörösére nőtt, az első négy év 3,5 százalékosával szemben az utolsó négy év átlagában már 12,4 százalékos volt.

5. *Lakáskarbantartás és egyéb kisebb szolgáltatások kiadásai.* Az igen heterogén kiadásokat magába foglaló csoport tételei három részre oszthatók, a lakással kapcsolatos kiadásokra, az egészségügyi tételeken kívüli személyi igények kielégítésére szolgáló tételekre és az egyéb kiadásokra (ruhatisztítás, cipőjavítás, híradástechnikai és konyhai eszközök javítása stb.).

A különböző karbantartásokra fordított kiadások aránya az összes költségen belül, más csoportoktól eltérően, nem mutat jelentős mozgást. A négyéves ciklusok közötti különbséget elsősorban az egy-egy alkalommal a lakással kapcsolatos nagyobb összeget igénylő tételek, mint lakásfestés, bútorok nagyobb javítása, kárpitozás cseréje vagy éppen ezek elmaradása okozza. Így a lakással kapcsolatos kiadások mutatják a legnagyobb ingadozást, az összes kiadáson belüli 3,1-5,8 százalék, de 2007-ben két százalék alatti aránnyal. A személyi tételekkel kapcsolatban az időszak nagyobb hányadában az összes kiadáson belüli 1,3 százalékról 2,4 százalékra emelkedés volt jellemző, a változatlan igények kielégítéséhez szükséges szolgáltatások fajlagos költségeinek viszonylag gyors emelkedése miatt, majd az utolsó ciklusban már az igények visszafogásával sikerült a kiadások részarányát ismét a korábbi szinthez közelíteni. Az egyéb tételek kezdeti magasabb arányát a kilencvenes évek közepétől nagyobb odafigyeléssel és a felmerülő munkákhoz az olcsóbb szolgáltatók megválasztásával, általában márkaszervizek elkerülésével sikerült a házaspárnak a korábbi 4 százalék körüli arányról annak mintegy felére csökkenteni.

Együttesen a karbantartásokra és egyéb kisebb szolgáltatásokra fordított kiadások aránya így nem nőtt, sőt a vizsgált első évekhez viszonyítva valamelyest csökkent is. Jelentősen alacsonyabb volt a kiadások aránya az utolsó évben, 2007-ben, amikor mint látni fogjuk, egy egyszeri beszerzés szükségessége háttérbe szorította a lakással kapcsolatban tervezett, de halasztható karbantartásokat

6. *Élvezeti cikkek.* Vitatható, hogy az élelmiszerek részbeni kiegészítésére szolgáló termékek, mint a kávé, a tea, a kakaó, korlátozott mennyiségben mennyiben sorolandók az élvezeti cikkek közé, míg a dohány- és alkoholtermékek általában egyértelműen annak tekinthetők.

A nem dohányzó házaspár esetében ma már a fogyasztás nagyjából havi 0,5 kilogramm kávé, 60 darab filteres teatasak, 5-10 dekagramm kakaó felhasználását és délben-este 1-1 pohár bor vagy helyette nyáron 1-1 pohár sör fogyasztását jelenti. Ez lényegesen kevesebb, mint volt az egyéb jövedelmet is biztosító és így a kiadásokat kevésbé érzékelő első évben, amikor még az összes kiadás 9,9 százalékát fordították élvezeti cikkekre, igaz akkor még több külföldi minőségi termék is bekerült a választékba. A teljes időszak kétharmadában általában 6 százalék körül volt az élvezeti cikkekre fordított kiadások aránya, az utolsó négy évben csökkent már 5 százalék alá az előbbieken felsorolt fogyasztással. Összességében a termékcsoport fogyasztásával kapcsolatban elmondható, hogy a kezdeti minőségi termékek fogyasztása fokozatosan átalakult kezdetben a kommersz, majd ma már

inkább a gyengébb minőségű termékek irányába és még a mennyiséget is egyre inkább korlátozza a házaspár.

7. *Ruházkodási és tartós fogyasztási cikkek vásárlása.* Ebbe a csoportba tartoznak mindazok a vásárlások, amelyek nem egyszeri, azonnali felhasználásra szolgáló, hanem akár a háztartásban, akár személyi használatban tartósan, hosszabb időn át használt javak beszerzését jelentik.

Amint a korábbi csoportoknál előfordulhat esetleg luxuskategóriába tartozó vagy minőségi színvonalú termékek vásárlása, úgy ebben a csoportban nyilvánvalóan sok esetben – különösen nyugdíjasoknál – olyan vásárlásra is sor kerül, ami az alapvető létszükségletű cikkek közé sorolható. Új cipő vagy kabát vásárlása nem tekinthető szabadon választott kiadásnak, ha az a korábbi, már használhatatlan vagy javíthatatlan darab pótlására szolgál. Ennek ellenére minden ilyen beszerzés is ebbe a csoportba került, mivel megvétele nem nélkülözhető ugyan, de általában halaszthatóbb, mint az élelmiszerek megvásárlása vagy a közüzemi díjak kifizetése.

A nélkülözhetőséget jól jellemzi a házaspár ilyen jellegű kiadása a még külön munkajövedelemmel is rendelkező első, majd a kizárólag a nyugdíjra támaszkodó további években. Az 1991 és 1994 közötti években havi átlagban 5 486 forintot költött a házaspár ilyen beszerzésekre, pótlásra, ami az összes kiadásuk 12,7 százaléka volt. Ez a két közbülső ciklusban 6,2 illetve 6,6 százalékra csökkent, majd az utolsó négy évben nominál értéken szinte megegyezett az első időszakkal, 5 653 forint volt havonként, az összes kiadáson belül 4,2 százalék. A részarány tehát kerekén 10 százalékponttal csökkent a 17 év alatt. (Megjegyezzük, hogy 2007-ben végleg javíthatatlanná vált a még 1989-ben vásárolt tévékészülék és annak pótlására került sor. Az új készülék sem nem plazma, sem nem LCD, hanem hagyományos képernyős készülék, amit emellett még „akcióban” vásárolt a házaspár, ennek ellenére a beszerzések részaránya a korábbi négy százalék körüli arányról éves átlagban 8,1 százalékra ugrott.)

Meg kell azonban azt is jegyezni, hogy a házaspár sok ilyen terméket kapott ajándékba gyermekeitől az évek folyamán. Ilyenek voltak karácsonyokra, születésnapokra adott ruhaneműk, cipők a feleségnek, technikai eszközök a férjnek, mint mobil telefon, elektromos írógép, majd lap-top, nyomtató stb., melyek alapvetően járultak hozzá a komfortérzet megtartásához, bizonyos területeken annak javulásához is, de mint az előbbieken láttuk a termékcsoport-kiadások közötti arányának mérsékelt szinten tartásához is.

8. *Családdal kapcsolatos kiadások.* A házaspárnak három gyermeke és kilenc, ma már felnőtt unokája van, valamint az első ciklus idején még élt a férj kizárólag özvegyi nyugdíjat élvező 80 év fölötti édesanyja. Az ebbe a csoportba tartozó kiadások ezekhez a személyekhez kapcsolódnak.

A külön háztartásban élő, a saját jogú minimál értékűnél kisebb összegű nyugdíjban részesült idős édesanya, 1993-ban bekövetkezett haláláig rendszeres támogatásra szorult. A gyerekeknél és az unokáknál ilyen igény nem merült fel, csak a szülői,

nagyszülői kapcsolatból fakadó megemlékezések költségei kerültek a kiadások közé. A gyerekek és házas társaik (6 személy), valamint az unokák (9 fő), összesen 15 személy, csak a karácsonyt, a születés- és névnapokat számítva 45 olyan alkalom van egy évben, amikor bármilyen szerény módon, de valamilyen ajándékkal, megvendéggeléssel számolni kell.

A kiadások összege az évek során természetesen jelentősen nőtt az árak emelkedésével párhuzamosan, az aránya azonban lényegében változatlanul az összes kiadás 11 százaléka körül alakult, lényegében változatlan maradt. Számottevő csökkentését az idős emberek minimális jóérzéséhez tartozó megnyilvánulások erőteljes korlátozása tenné csak lehetővé, amire a házaspár eddig nem gondolt és a jövőben is legfeljebb reáljövedelmének komoly visszaesése esetén volna hajlandó, még ha a 2007. évi 8,7 százalékos arány már erre is utalna.

9. Kultúra, reprezentáció. Ebbe a csoportba soroltuk értelemszerűen a hangverseny-, opera-, színház-, mozilátogatások, a könyv- és folyóirat-vásárlás, a múzeumlátogatás, egy-egy rövid külföldi út költségeit, valamint a nem családtagok vendéglátásával kapcsolatos kiadásokat. Ezek a szellemi színvonal meg- és fenntartásához is nélkülözhetetlen kiadások kifejezetten csökkentek reálértékükben. Az időszak első ciklusában még 10 százalékos meghaladó arányuk már a második ciklusban 8,2 százalékra esett vissza, majd tovább mérséklődve 7 százalék fölé stabilizálódott, majd egy rövid berlini út költségei miatt 2007-ben 9,4 százalék lett.

A korukkal járó kedvezmények a múzeumok állandó és egyes időszaki kiállításainak látogatását még lehetővé tették, de már elmaradtak az opera- és hangversenylátogatások, amiket a rádió és televízió adásai pótolnak, évi egy-két esetre korlátozódott a mozilátogatás, megszüntették a folyóirat- és újságvásárlást, a hiányt a világhálóra kapcsolva igyekeztek mérsékelni és minimálisra csökkentették (csak a számukra tartamilag legértékesebbre korlátozták) a korábban szabadabban megengedett könyvvásárlást. Ugyancsak visszafogták az amúgy sem széles körű társas életüket is, lényegesen mérsékelve az ezzel járó reprezentációs kiadásokat.

10. Közlekedés, gépkocsi fenntartása. A házaspár egyetlen valóban luxuskiadása volt (amennyiben annak lehet számítani), gépkocsijuk fenntartása, üzemeltetése. A férj nyugdíjba vonulásakor új Polski Fiat 126 típusú kis kocsit vettek, amit később egyik fiúk már 75 ezer kilométert futott kis Renault 5 típusú kocsijára cseréltek. A kocsit használata utolsó éveiben egyik távolabb lakó gyermekük látogatására, a kéthetenkénti bevásárlás lebonyolítására, valamint kirándulásokra, átlagosan havi 500 kilométer futásra korlátozták. Az utolsó évek „arany szabálya” volt, hogy legfeljebb, de nem rendszeresen, havonként engedtek meg egy tankolást (40 liter benzin).

Az erős korlátozás eredménye volt, hogy az időszak első felében még átlagosan a kiadások 12 százaléka körüli hányadát fordították közlekedésre, ez a harmadik ciklusban már csak 8 százalék körül alakult. Az általános drágulás, ezen belül is az árarányok és így a kiadásokon belül a kiadások arányainak megváltozása arra készítette

a házaspárt, hogy 2004-ben megváljon „luxuseszközétől”, aminek eredményeként az utolsó ciklusban a közlekedésre fordított kiadásaik aránya 3,5 százalékra csökkent. A már említett néhány napos berlini út „fapados repülőjáratának” költsége 2007-ben ezt az arányt ismét 6,1 százalékra emelte.

11. Egyéb kiadások. Mint minden pénzügyi elszámolásnál, itt is szerepel az egyéb rovat, amelybe az előző csoportokba nem sorolható tételek vagy éppen kisebb kiadások kerültek, melyeket a pontos megnevezés hiányában más csoportba nem lehetett sorolni. Ezeknek a kiadásoknak az aránya egy és két százalék között alakult a ciklusok során. Nincs szó tehát lényeges tételekről, csak a rendelkezésre álló források és az elszámolt kiadások egyenlőségét biztosítja ez a kiadási csoport.

A részleteket összefoglalva a kiadások összetételében jellegzetes változások történtek az elmúlt tizennyolc év során, amelyek kisebb-nagyobb hasonlósággal minden nyugdíjas háztartásban érvényesültek. Ezek a változások a következők:

– alapvetően megváltoztak a létfenntartási és a szabadon választott kiadási tételek arányai, az előbbiek az időszak kezdetének 45 százalék körüli arányáról az összes kiadások kétharmadára nőttek és ennek megfelelően csökkentek a szabadon választott kiadások 55 százalékáról a kiadások egyharmadára. A nyugdíjon kívüli munkajövedelem ilyen szempontból torzító hatását figyelmen kívül hagyva is jelentős volt a változás, mivel a második ciklus óta is tíz százalékpont volt a változás a csoportokban, így a létfenntartási kiadások 56 százalékáról 66 százalékra emelkedtek és értelemszerűen csökkentek a szabadon választott kiadások,

– három alig befolyásolható termékcsoporthoz, a fűtés- háztartási energia, a lakással kapcsolatos állandó költségek és az egészségügyi kiadások együttes aránya a második ciklus 29 százalékáról a negyedik ciklusra 44 százalékra emelkedett, így a nyugdíjak reálértékének lényegében változatlan szintje miatt takarékosagra csak két másik tételnél, az élelmiszer és háztartási vegyi áru és a karbantartási kiadások területén volt lehetőség, amivel élt is a házaspár e két tétel arányát 27 százalékról 20 százalék körülire csökkentette az összes kiadásokon belül,

– a kiadások viszonylag egyenletes, fokozatos csökkenés ment végbe az évek során, engedve a létfenntartási kiadások nyomásának, amit minden területen az igények visszaszorításával és a már amúgy is mérsékelt gépkocsi-használatuk megszüntetésével igyekeztek korlátok közé szorítani.

A házaspárnak a nyugdíjasok átlagánál magasabb jövedelme korábban a létfenntartási és a szabadon választott kiadások közötti arány igen kedvező szintjét biztosí-

totta. Az ellátásnak a megfigyelt utolsó évekre végeredményben változatlan reálértéke, ugyanakkor egyes költségcsoportokban az átlagost lényegesen meghaladó ár-emelkedések a két kiadáscsoport között igen nagy változást okoztak, amivel a nyugdíjasok minden rétegének számolnia kellett. Különbség csak annyi lehet, hogy amit a házaspár a maga szintjén még „belső átcsoportosításokkal” meg tudott oldani, az átlagos vagy a még azt sem elérő nyugdíjat „élvezőknél” már a szegénységi küszöb, akár a létminimum alá szoríthatta az életszínvonalat.

3. Tizennyolc év mérlege

A több mint másfél évtized alatt végbement változások egybevetése előtt az elemzőnek kötelessége bemutatni a választott kiinduló évnek, az 1990-nek jellemzőjét, azt hogy az a vizsgálat nézőpontjából gyenge, átlagos, netán kedvező esztendő volt-e. A KSH adatai szerint a tíz évvel korábbihoz, 1980-hoz képest választott bázisévben, 1990-ben a nemzetgazdaság bruttó hazai terméke (GDP) 12,0 százalékkal, a lakosság jövedelmének reálértéke 13,3 százalékkal, az egy főre jutó fogyasztás reálértéke 10,2 százalékkal volt több, az egy ellátottra jutó nyugdíj reálértéke pedig 6,1 százalékkal volt magasabb, mint tíz évvel előtte. Vagyis a növekedési mutatók évi átlaga gyenge 1 százalék körüli, a nyugdíjak évi átlagos emelkedése pedig ezek mintegy fele, 0,5 százalék körüli volt, így a nyugdíjasok nézőpontjából a választott bázisév a megelőző tíz esztendő negatívumait tükröző kedvezőtlen esztendő volt, az innen történő minden pozitív irányú változás vitathatatlanul alacsony szintről történő elmozdulás. Ezt mindenképpen figyelembe kell venni az elért eredmények értékelése során.

A dolgozat nem foglalkozik a nyugdíjrendszer sokat vitatott kérdéseivel, annak fenntarthatóságával, egy-két adat az egybevetés érdekében mégis ide kíváncsodik. Az Országos Nyugdíjbiztosítási Főigazgatóság adatai szerint a ma 3 millió nyugdíjas-ként említett társadalmi rétegben élők száma ebben az összetételben 1990-ben 2,6 millió fő volt, vagyis nincs szó megkétszereződésről. Az öregségi nyugdíjasok száma 1 462 ezer főről, 2006-ban 1 665 ezer főre gyarapodott. Ezek átlagos havi nyugdíja 6 450 forintról, a házaspár nyugdíjának növekedéséhez hasonlóan, nőtt és ért el 2006-ra 69 145 forintot. Nem hagyható figyelmen kívül, hogy ez utóbbi összegben belül a korbetöltött öregségi nyugdíjasok ellátásának átlagos összege 63 759 forint, az előrehozott nyugdíjaké 90 552 forint, a karkedvezményes nyugdíjaké pedig 110 444 forint volt e két utóbbi csoport 200 ezer főt kitevő rétegénél. A klasszikus értelemben vett, tehát korbetöltött öregségi nyugdíj reálértéke a másfél millió nyugdíjasnál 2007-ben ugyanúgy az 1990. évi szinten mozgott, mint a házaspáré.

A 2007. évi nyugellátás reálértéke a feleségnek a női átlagos nyugdíjnak megfelelő színvonalú ellátásában 18 év után 10 százalékkal volt magasabb, mint 1990-ben, míg a férjnek az átlaghoz viszonyítva magasnak tekinthető nyugdíja még 2007-ben is 4,4 százalékkal maradt el reálértékben a bázisévi szinttől. A házaspár együttes nyugdíja így, mintegy 15 évnyi visszaesés után, az utolsó két évben érte el ismét az 1990. évi színvonalat. Az együttesen is az átlagot meghaladó nyugellátás a bázis évben a nemzetgazdaságban foglalkoztatott szellemi dolgozók nettó keresetének még 82,3 százaléka, 2006-ban már csak 65,3 százaléka volt. A visszaesést az okozta, hogy egyidejűleg a nemzetgazdaságban foglalkoztatottak átlagos nettó keresetének reálértéke összesen 28,0 százalék emelkedést ért el. Megjegyzendő, hogy egyidejűleg a változatlan árakon számított bruttó hazai termék (GDP), az ország teljesítménye 38,5 százalékkal nőtt. Annak ellenére, hogy a dolgozat nem foglalkozik a nyugdíjrendszerrel kapcsolatos leginkább vitát kiváltó kérdésekkel, az viszont vizsgálatai alapján egyértelműen leszögezhető, hogy a már 1990-ben nyugdíjas idős emberek ellátásának reálértéke hosszú évek jelentős visszaesése, majd mintegy tíz évnyi növekedés után 2007-re még éppen csak elérte az 1990. évi reálértékét, szemben a keresetek és a GDP viszonylag jelentős növekedésével. Az 1997. évi nyugdíjtörvény alapján a svájci indexálás segítségével a következő években érhetnének el szerény mértékű, a termelés és a keresetek növekedésének töredékét jelentő életszínvonal-javulást. Ez utóbbinak azonban továbbra is feltétele a kiadási tételeken belül az árak és árarányok változásának mikéntje. Az elmúlt tizennyolc évben a házaspár esetében a létfenntartási kiadások aránya az összes kiadáson belül 40 százalékról 67 százalékra nőtt. Az eddigiek során viszonylag kismértékű változás történt az élelmiszerek és háztartási vegyiárak csoportjánál, az ide tartozó termékek árszínvonalának emelkedése eddig nem érte el az áremelkedés átlagos mértékét. Kifejezetten csökkent a karbantartási kiadások aránya, ami azonban már kevésbé az árak alakulásának, inkább az árak és szolgáltatások igénybevételének mérsékléséből adódott. A létfenntartási kiadások növekedésének döntő részét, a változatlan mennyiségi felhasználás ellenére, a fűtés, háztartási energia és a lakással kapcsolatos állandó költségek arányának együttesen az összes kiadáson belül 17 százalékról 30 százalékra emelkedése adta. Egyrészt az öregedés következtében természetesen jelentkező növekvő igény az egészségügyi ellátás iránt, másrészt az ezen a területen, elsősorban a gyógyszereknél tapasztalt ugyancsak az átlagot lényegesen meghaladó áremelkedés okozta az egészségügyi kiadások arányának is a jelentős emelkedését a kiadások között.

A szabadon választott kiadások közé sorolt csoportok között, az előbbiekből adódóan, növekvő arányokkal nem lehetett számolni, a kérdés csak az, hogy hol milyen mértékű visszafogás sújtja, vagy legalábbis zavarja a nyugodt életet. A házaspár esetében minden kiadáscsoportban nagyjából azonos mértékű visszafogás történt, jelentős mértékű csökkenést a közlekedés, gépkocsifenntartás tételeknél tudtak elérni a használat kezdeti korlátozásával majd a gépkocsitartás felszámolásával. Sok tételnél

ezekben a csoportokban nem lehet megállapítani azt a „mennyiségi normát”, ami még vagy már kielégítőnek tekinthető a nyugdíjas háztartásban, de akár a múlttal való összehasonlítás, akár az általános életszínvonalhoz társított egybevetés azt mutatja, hogy a házaspár szabadon választott kiadásai általánosságban elmaradnak nemcsak saját igényüktől, de az általában elvárhatótól is.

A jövőbe tekintve a nyugdíjak növekedése az életszínvonal emelkedésének lehetőségét vetíti előre, amennyiben nemzetgazdasági érdekekre hivatkozó politikai döntés nem véli a már elért színvonalat is finanszírozhatatlannak és mint az a vizsgált időszakban már megtörtént, a nyugdíjakra fordított kiadások visszafogásának keretében az öregségi nyugdíjak reálértékének „mérséklésével” kívánja biztosítani a nyugdíjasok hozzájárulását is a nemzetgazdaság egyensúlyának fenntartásához. A dolgozatban arra a megállapításra jutottunk, hogy a házaspár, de általában az öregségi nyugdíjat élvezők 2006-ban érték el ismét az 1990. évi ellátásuk reálértékének szintjét. A KSH adatai szerint 1990-ben, amikor az öregségi nyugdíjasok csoportjában a létminimum havi 4 998 forint, az öregségi nyugdíjak átlagos összege 6 450 forint volt, a létminimumot 29,1 százalékkal haladta meg, ami arra utal, hogy a másfél millió főnyi korbetöltött nyugdíjas nem kis hányada élt a létminimum alatt. Ez az arány 1995-ig jelentősen romlott, mivel ebben az évben a 13 519 forint összegű átlagos öregségi nyugdíjjal szemben a létminimum értéke 14 792 forint volt, 9,4 százalékkal több, mint a nyugdíj, vagyis a létminimum alatt élni kényszerülők voltak többségben. Öt évvel később, 2000-ben a 31 028 forint összegű létminimumot már 7,2 százalékkal meghaladta az öregségi nyugdíj átlagos 33 258 forint összege és 2006-ban az 54 115 forint összegű létminimumhoz képest az 1990. évhez hasonlóan 27,8 százalékkal volt több a 69 145 forint összegű átlagos öregségi nyugdíj. Hozzá kell azonban ehhez tenni, hogy amint a házaspár kiadásainak változásánál láttuk, az egész nyugdíjas rétegre is jellemző, hogy a létfenntartási költségek aránya lényegesen megemelkedett a kiadásokon belül, jelentősen csökkentve a szabadon választott kiadások arányát, így mérsékelve azt a hatást, amit egyébként az ellátások kétségtelenül növekedő reálértéke kifejtett volna. Az élelmiszerek és az energiahordozók árának az átlaghoz képest várhatóan további erőteljesebb növekedése ezt a tendenciát a nyugdíjak reálértékének remélt növekedése ellenére, sajnos, várhatóan továbbra is fenn fogja tartani.

Irodalom- és forrásjegyzék

ARTNER A. – INOTAI A. [1997]: Felzárkózási esélyek a statisztikai adatok alapján. *Statisztikai Szemle*. 75. évf. 4–5. sz. 293–302. old.

EUROSTAT [2007]: *ECU/EUR exchange rates versus national currencies*. Statistical Office of the European Communities. Brussels.

- HÁGEN I. ZS. – KIRÁLY ZS. [2004]: *A jelenlegi európai uniós tagországok és az unióhoz csatlakozó közép-kelet európai országok összehasonlító vizsgálata néhány mutató alapján*. Károly Róbert Főiskola. Gyöngyös.
- KSH [2003]: *Magyar Statisztikai Évkönyv, 2002*. Központi Statisztikai Hivatal. Budapest.
- KSH [2004]: *Időskorúak Magyarországon*. Központi Statisztikai Hivatal. Budapest.
- KSH [2005]: *A jövedelem mint az anyagi jólét és a szegénység mérőszáma*. Társadalomstatistikai Füzetek, 43. Központi Statisztikai Hivatal.
- KSH [2006]: *Létminimum 2005*. Központi Statisztikai Hivatal. Budapest.
- KSH [2007]: *Makrogazdaság, 2006*. Központi Statisztikai Hivatal. Budapest.
- KSH [2008]: stADAT-táblák. 6. Devizaárfolyamok. 7. Nemzetközi statisztika. Központi Statisztikai Hivatal. Budapest.
- KSH [2008]: *Létszám és kereset a nemzetgazdaságban, 2007*. január-november: Gyorstájékoztató. 12. sz. Január 17. Központi Statisztikai Hivatal. Budapest.
- KSH [2008]: *A háztartások jövedelmének és fogyasztásának színvonala és szerkezete*. Statisztikai tükör. II. évf. 1. sz. Január 14. Központi Statisztikai Hivatal. Budapest.
- KSH [2008]: *Létszám és kereset a nemzetgazdaságban*. Gyorstájékoztató 30. sz. Február 20. Központi Statisztikai Hivatal. Budapest.
- ONYI [2008]: *Nyugdíjban és nyugdíjszerű ellátásban részesülők száma és a teljes ellátás átlagos összege évenként (1990–2007)*. Országos Nyugdíjfolyósító Intézet. Budapest.

Summary

The author presents the change of financial position of senior citizens through the example of a couple already retired in 1990. The real value of pensioners' allowances decreased steadily in the first six years of the above period – except one year of election – so that, in 1996 it was less than 70 percent of the starting value. Over the period of the following ten years a continuous growth characterised the changes, still in 2006/2007 the real value of pensioners' incomes met only the level of 1990. The increase was differentiated: while the real value of the husbands' above average pension did not achieve in 2007 the level of 1990, the wife's lower allowance of average level, as a result of several minor corrections, exceeded the starting value by 10 percent in 2007 compared to 1990. Major changes characterised the structure of expenses. The share of three elements of the cost of living (heating, flat, health) grew dramatically reaching 44 percent by 2007, while in 1990 it covered only 18 percent. As a result, the amount of disposable income dropped considerably. Based on statistical data, the above trend is effective in case of 1,6 million people living on pensioners' allowances. However, the actual regulation of pensions aims at a moderate increase of real value, the prices of food and energy will most probably exceed this growth, resulting, the ratio of optional expenses (culture, transport, consumer goods) will decline within the budget of senior citizens.

Hiányzó adatok és kezelésük a statisztikai elemzésekben

Oravecz Beatrix,

a Budapesti Corvinus Egyetem
tanársegédje

E-mail: beatrix.oravecz@uni-corvinus.hu

Adathiánnyal szinte minden adatbázis esetén találkozunk. A hiányzó adatokat valamilyen módon kezelni kell, nem hagyhatjuk ki őket egyszerűen a mintából, mert a sokasági paraméterbecslések torzítottak lehetnek, hacsak az adathiány nem teljesen véletlenszerű. A hiányzó adatok kezelésének célja éppen ennek a torzításnak az eltüntetése. Ezt a célt a különböző módszerek annak függvényében érik el, hogy mennyire helyesen sikerül azonosítani és modellezni az adathiány sajátosságait. Ebben a tanulmányban áttekintjük a hiányzó adatok típusait és a kezelésük lehetséges módjait, kiemelve az egyes módszerek előnyeit, hátrányait és alkalmazásuk feltételeit. A hiányzó adatok kezelésére nem létezik egyetemesen legjobb megoldás. Lényeges szempont, hogy a választott eljárás összhangban legyen a később elvégzendő elemzésekkel, és az olyan adatbázisok esetében, ahol a hiányzó adatokat valamilyen módon pótolták, a felhasználók is láthassák az adatpótláshoz használt módszert.

TÁRGYSZÓ:

Statisztikai mintavétel.
Statisztikai módszer.
Statisztikai elemzés.

A hiányzó adatok sok kutatásnál okoznak problémát, mert a minta véletlenszerűségét rombolhatják le, pedig a legtöbb statisztikai módszer és következtetés alapja a véletlen minta. Ebben a tanulmányban röviden áttekintjük a hiányzó adatok típusait és a kezelésükre használt legelterjedtebb módszereket, kiemelve fő előnyeiket és hátrányaikat.

Egy általános adatmátrix sorai tartalmazzák a megfigyelési egységeket, vagy eseteket, az oszlopok pedig a változókat, amelyek értékét minden egység esetén ismerjük. Az adatmátrixban lévő adatok általában valós számok, amelyek vagy egy mennyiségi ismérv tényleges értékeit fejezik ki (például az életkor vagy a jövedelem), vagy egy minőségi ismérv kategóriáit reprezentálják (például az iskolai végzettség vagy a nem). A gyakorlatban azonban az a jellemző, hogy ez az adatmátrix nem teljes, bizonyos értékek hiányoznak.

Például egy háztartási bevételeket és kiadásokat vizsgáló kutatás során a megkérdezettek megtagadhatják a jövedelemre vonatkozó kérdés megválaszolását, vagy egy fogyasztói preferenciákat vizsgáló kutatás során előfordulhat, hogy a válaszadó nem tud választani két termék közül, egyiket sem preferálja a másikkal szemben. Az első esetben a jövedelem értékét tekinthetjük hiányzónak, hiszen van mögötte egy tényleges érték, csak mi nem ismerjük. A második esetben azonban nem tekinthetjük a termékpreferenciát hiányzónak, mert nincs mögötte valós érték, a válaszadó nem megtagadta a választ, hanem nem tudott válaszolni. Ebben az esetben a „nincs preferencia” vagy „nem tudom” is egy válaszadói réteget jelöl. A legtöbb statisztikai szoftver tartalmaz egy vagy több speciális kódot az adathiány bevitelére. Egynél több kód lehetővé teszi a különböző jellegű adathiányok beazonosítását, mint „nem tudja”, „válasz megtagadás”, „értelmetlen adat”. Ez utóbbi esetben van ugyan adatunk, tehát látszólag nincs adathiány, de tudjuk, hogy az nem megbízható, vélhetően hibás, így azt valójában nem használhatjuk az elemzésekben, hanem a hiányzó adatokhoz hasonlóan kell kezelniük¹. Felmerülhet a kérdés, miért kell egyáltalán a hiányzó adatokkal foglalkozni, ahelyett, hogy egyszerűen törölnénk őket a mintából. Válaszként álljon itt a következő példa.

1992. április 9-én a Konzervatív Párt megnyerte a brit választásokat, ami óriási bukást jelentett a közvélemény-kutatási iparágnak. A választások napján a négy legnagyobb közvélemény-kutató cég a Munkáspárt 0,9 százalékos pontos győzelmét várta.

¹ Az outlierek esetében is van adatunk, de azt nem célszerű a többihez hasonló módon használni. Az outlierek meghatározásáról és kezeléséről olvashatunk például *Csereháti* [2004] cikkében. Ez utóbbi típusú „adathiányok” kezelésében az adatellenőrzésnek és korrekciónak nagy szerepe van, de ezekkel ebben a tanulmányban nem foglalkozunk.

Ezzel szemben a Konzervatív Párt győzött 7,6 százalékponttal. Ez 8,5 százalékpontos hiba, ami igen nagy. Egy utólagos vizsgálat megállapította, hogy a hiba fő oka az volt, hogy a kutatás során nem foglalkoztak a válaszmegtagadásokkal és a „még nem tudom” típusú válaszokkal, hanem egyszerűen törölték őket a mintából. Ez a gyakorlat végzetes volt az eredmények szempontjából, mert az utólagos kutatás megmutatta, hogy a konzervatív pártiak kevésbé tárták fel választási szándékukat. (Hasonló volt a helyzet a magyarországi 2002-es választások során is.)

Látható tehát, hogy a hiányos adatbázisokból való következtetések torz képet adhatnak. Törekedni kell tehát az adathiány természetének megismerésére, majd ezen információk figyelembevételével a hiányzó adatok valamilyen kezelésére.

A tanulmányban először áttekintjük az adathiányok jellemző mintázatait, majd megvizsgáljuk az ún. adathiány-mechanizmusokat, végül sorra vesszük azokat a lehetséges eljárásokat, amelyek adathiányos helyzetekben alkalmazhatók.

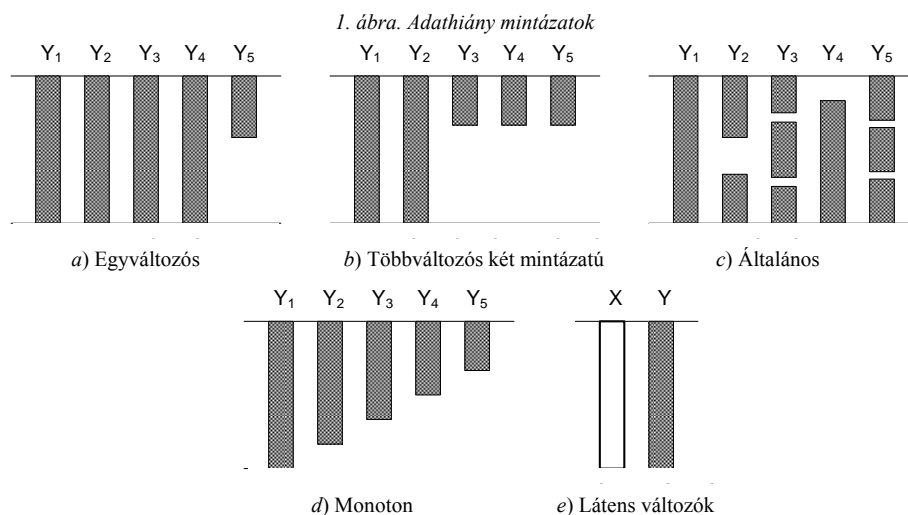
A hiányzó adatok kezelésével foglalkozott korábban a *Statisztikai Szemlében* Máder Miklós Péter [2005] „Az imputálási eljárások hatékonysága” című cikke. Ez a korábbi cikk nem foglalkozott az adathiány mintázatokkal, ezért ezeket ebben a tanulmányban ismertetjük. Máder cikke néhány eljárás hatékonyságát vizsgálta modellezéssel. Ez a tanulmány nem tartalmaz empirikus vizsgálatot, hanem az ott alkalmazott és egyéb alkalmazható módszerek elméleti hátterét és tulajdonságait tekinti át.

1. Hiányzó adatok típusai

A következőkben áttekintjük az adathiány típusait. A csoportosítás egyik szempontja az *adathiány mintázata*. A mintázat azt írja le, hogy mely adatok a megfigyelték és mely adatok hiányoznak az adatmátrixban. A másik csoportosítási szempont az *adathiány-mechanizmus*, amely a hiányzás és az adatbázisban szereplő változók értékei közötti kapcsolatot veszi figyelembe.

1.1. Adathiány mintázat

Legyen $\mathbf{Y} = (y_{ij})$ egy $(n \times K)$ általános adatmátrix, hiányzó adatok nélkül, amelynek i -dik sora $y_i = (y_{i1}, \dots, y_{iK})$, ahol y_{ij} az Y_j változó értéke az i -dik egységnél. Hiányzó adatok esetén legyen $\mathbf{M} = (m_{ij})$ az adathiány indikátor mátrix (*Little–Rubin* [2002]), ahol $m_{ij} = 1$, ha y_{ij} hiányzik és $m_{ij} = 0$, ha y_{ij} megfigyelt. Az \mathbf{M} mátrix definiálja az adathiány mintázatot. Az 1. ábra mutat néhány példát az adathiány-mintázatokra. (A megfigyelt y -ok ($m = 0$) sötétrel jelölve.)



Az *egyváltozós adathiány* az 1. ábra a) esete, amikor csak egyetlen változóban van adathiány, a többi változó teljesen megfigyelt. Ilyen mintázata lehet például a mezőgazdasági kontrollált kísérletek eredményének, ahol azt vizsgálhatják, hogy milyen a kapcsolat az Y_K eredményváltozó (terméshozam) és az Y_1, \dots, Y_{K-1} magyarázóváltozók (öntözővíz, hőmérséklet, műtrágya típusa, mennyisége) között. A magyarázóváltozók ekkor teljesen megfigyelték, nincs hiányzó adat, a függő változóban viszont előfordulhat adathiány (például hibás vetőmag vagy rossz adatrögzítés miatt).

A *többváltozós kétmintázatú* adathiány egy másik általános mintázat, amikor az előző példában szereplő egyetlen adathiányos változó (Y_K) helyett több adathiányos változónk van (Y_{j+1}, \dots, Y_K), ahol mindegyik egyformán megfigyelt, vagy hiányzik ugyanazokra az esetekre. (Lásd az 1. ábra b) esetét, ahol $K = 5$ és $J = 2$.)

Erre a mintázatra lehet példa a kérdőíves felméréseknél az egységintű nemválaszolás. (Amennyiben az adathalmazból egy-egy elem teljesen hiányzik teljes (vagy egységintű) nemválaszolásról (unit nonresponse) beszélünk.) Ez az egységintű nemválaszolás előfordulhat azért, mert a kiküldött kérdőívet meg sem kapta a címzett, vagy megkapta, de megtagadta a válaszadást. Ekkor a kérdőívben szereplő változók lesznek az adathiányos változók. A teljes, adathiányt nem tartalmazó változók a minta tervezéséhez használt változók lesznek, amelyek mind a válaszolók, mind a nemválaszolók esetében előzetesen ismertek egy listáról (például név \rightarrow nem, lakcím).

Általános adathiány-mintázat úgy alakul ki, ha csak bizonyos kérdésekre adott válaszok hiányoznak, ekkor részleges (vagy tétel szintű) nemválaszolásról (item nonresponse) beszélünk. Ebben az esetben az adathiány mintázata általában semmi-féle specialitással nem rendelkezik. (Lásd 1. ábra c) esetét.)

Monoton adathiány következik be például, ha a longitudinális felmérések időről időre gyűjtenek be adatokat ugyanazon megfigyelési egységekről. Ezekben a felmérésekben gyakori jelenség a lemorzsolódás, ami azt jelenti, hogy a megfigyelési egység kiesik a mintából, még a kutatás befejezése előtt. Például háztartás-panel esetén a család külföldre költözik, vagy klinikai kísérleteknél más gyógyszerek hatása, vagy egyéb betegség miatt a beteg nem tud tovább részt venni a kísérletekben. A lemorzsolódás egy példája a monoton mintázatú adathiányoknak. (Lásd 1. ábra *d*) esetét.) Ekkor a változókat lehet úgy sorba rendezni, hogy minden Y_{j+1}, \dots, Y_K hiányzik, ha Y_j hiányzik. Vannak olyan módszerek, amelyek csak az ilyen mintázatú adathiányt tudják kezelni. Az ilyen mintázat a gyakorlatban ritkán fordul elő, közel monoton mintázat azonban már gyakrabban.

A nem megfigyelhető *látens változókat* is felfoghatjuk adathiány problémaként, csak ezeknél a látens változóknál speciálisan minden megfigyelési érték hiányzik. Az 1. ábra *e*) esetében az X jelenti a látens változók csoportját, ahol minden érték hiányzik és Y pedig a teljesen megfigyelt változók csoportját. Ekkor természetesen bármiféle elemzéshez különböző feltételezésekkel kell élnünk. Látens változó lehet például a klinikai kísérleteknél a beteg gyógyulásba vetett hite, ha erre vonatkozóan nem szerepelnek adatok a mintában.

1.2. Adathiány-mechanizmus

A hiányzó adatok kezelésének legalkalmasabb módját akkor tudjuk megtalálni, ha ismerjük, hogy miként lettek hiányzóak. *Little* és *Rubin* [1987] az adathiány három alapvető esetét különbözteti meg, attól függően, hogy milyen a kapcsolat a hiányzás és az adatbázisban levő változók értékei között. Ezeket ők *adathiány-mechanizmusnak* nevezték el.

Intuitíve és formálisan is megadjuk az egyes csoportok definícióját. Legyen továbbra is az $\mathbf{Y} = (y_{ij})$ a teljes adatmátrix és az $\mathbf{M} = (m_{ij})$ az adathiány indikátor mátrix. Az adathiány mechanizmus jellemezhető az \mathbf{M} adott \mathbf{Y} melletti feltételes eloszlásával, az $f(\mathbf{M}|\mathbf{Y}, \theta)$ -val, ahol θ ismeretlen paramétereket jelöl.

A *teljesen véletlenszerű adathiány* (Missing Completely at Random – MCAR) esetében a teljes adatállománnyal rendelkező egységek és a hiányzó adatokat tartalmazó egységek teljesen egyformák, ugyanazon eloszlásból származnak.

A hiányzás tehát nem függ az \mathbf{Y} értékétől, sem a megfigyelt, sem a hiányzó adatokkal rendelkező változók értékétől, azaz:

$$f(\mathbf{M}|\mathbf{Y}, \theta) = f(\mathbf{M}|\theta), \quad \text{minden } \mathbf{Y}, \theta \text{ esetén.} \quad /1/$$

Ez a mechanizmus például akkor fordulhat elő, ha minden válaszadó egy pénzérmé feldobásával dönti el, hogy válaszol-e a kérdésre.

Véletlenszerű adathiány (Missing at Random – MAR) esetében a hiányzó adatokat tartalmazó egységek eltérnek a hiánytalan adatokkal bíró egységektől, de a hiány jellegzetességei nyomon követhetők, előre jelezhetők az adatbázis más változói segítségével. Az adathiány tehát más változókkal kapcsolatban van, de azzal a változóval, amelyikben a hiányzás felmerül nincs közvetlen kapcsolatban.

Legyen $Y_{megfigyelt}$ azon változók halmaza az Y -ből, amelyben nincs adathiány és $Y_{hiányzó}$ azon változók halmaza, amelyben van adathiány. A véletlenszerű adathiány tehát az jelenti, hogy:

$$f(M|Y, \theta) = f(M| Y_{megfigyelt}, \theta), \quad \text{minden } Y_{hiányzó}, \theta \text{ esetén.} \quad /2/$$

Ez a mechanizmus fordul elő például, ha a magasabb jövedelemmel rendelkezők nagyobb valószínűséggel tagadják meg a jövedelemre vonatkozó kérdések megválaszolását, de a jövedelemre következtetni tudunk a felmérés más változói (például: fogyasztási szokások, fogyasztás és megtakarítás egymáshoz való viszonya) alapján.

A *nem véletlenszerű adathiány* (Not Missing at Random – NMAR vagy másként „nonignorable”, nem elhanyagolható) esetében az adathiány nem véletlenszerű, és más változókkal sem becsülhető, mert közvetlenül az adathiányt tartalmazó változóval van kapcsolatban. Az M eloszlása tehát függ az Y hiányzó értékeitől (is). Ez az adathiány legveszélyesebb, legnehezebben kezelhető formája.

Ez a mechanizmus fordul elő például, ha a magasabb jövedelemmel rendelkezők nagyobb valószínűséggel tagadják meg a jövedelemre vonatkozó kérdések megválaszolását, és a jövedelemre nem tudunk következtetni a felmérés más változóiból.

A hiányzó adatok számos problémát okoznak. Ugyanazon az adatbázison különböző kutatók által végzett elemzések eredménye között inkonzisztenciát tapasztalhatunk, ha azok másképpen kezelték a hiányzó adatokat. A hiányzó adatok kezelésére pedig azért van szükség, mert a sokasági paraméterbecslések torzítottak lehetnek (mint ahogy az 1992-es brit választásoknál is történt), ha csak az adathiány nem teljesen véletlenszerű.

A hiányzó adatok kezelésének célja éppen ennek a torzításnak az eltüntetése. Ezt a célt a különböző módszerek annak függvényében érik el, hogy mennyire helyesen sikerül azonosítani és modellezni az adathiány sajátosságait.

2. Hiányzó adatok kezelésére szolgáló módszerek

A hiányzó adatokkal való elemzés irodalma nem túl hosszú múltra tekint vissza. A szakirodalomban ajánlott és alkalmazott módszereket a következőképpen csoportosíthatjuk (Little–Rubin [2002]).

1. Teljesen megfigyelt vagy elérhető egységek elemzésén alapuló eljárások
2. Átsúlyozás

3. Imputációalapú eljárások

4. Modellalapú eljárások

A csoportok nem átfedésmentesek, de ebben a csoportosításban tekintjük át az alábbiakban a nemválaszolások kezelésének legelterjedtebb módszereit. A felsorolás nem tartalmaz minden alkalmazható módszert, csak a széles körben használt megközelítéseket.

2.1. Teljesen megfigyelt vagy elérhető egységek elemzésén alapuló eljárások

Az *adathiányt tartalmazó esetek törlését* (listwise vagy casewise adat törlés) említjük elsőként. Ha egy megfigyelési egységnél akár csak egy változó tekintetében is hiányzik adat, az egész megfigyelést törlik az adatbázisból. Az eljárást számos statisztikai programcsomag tartalmazza alapmegoldásként. A megoldás előnye az egyszerűsége, és az hogy az egyváltozós statisztikák összehasonlíthatók, mert mindegyik ugyanazon adatokon alapulva lett számítva. Hátránya viszont, hogy a nem teljes megfigyelésekben meglévő információt egyáltalán nem hasznosítja. Csak teljesen véletlenszerű eredetű adathiány esetén alkalmazható, azaz ha a hiányzó adatokat tartalmazó esetek az összes eseten belüli véletlenszerű almintának tekinthetők. Ha az adathiány nem MCAR, akkor a módszer torzítást okoz. Relatív alacsony nemválaszolási arány mellett ésszerű lehet az alkalmazása, mert ekkor az egyszerűségből fakadó előnyök ellensúlyozhatják a néhány hiányzó adat által okozott információvesztést és minimális torzítást.

Az *elérhető adatok elemzése* (available case analysis) a második módszer. Az adatok törléséből származó információvesztés csökkenthető, ha minden változó elemzésekor az abban a változóban meglévő összes adatot használjuk. A módszer hátránya, hogy a változónkénti elemzések más-más adatbázison készülnek, így az eredmények összehasonlítása problémás lehet. E módszer alkalmazásakor kétváltozós korreláció- vagy kovariancia-számításhoz mindig az adott két változó tekintetében elérhető adatpárokat használják (pairwise available case). Számos statisztikai programcsomag tartalmazza ezt a kezelési módot. Előnye, hogy jobban kihasználja a meglévő adatokat, de az eredményeként létrejövő korrelációs mátrix nem feltétlen lesz pozitív definit. (Ekkor pedig ez a mátrix már nem is nevezhető korrelációs mátrixnak.)

Nézzük a következő példát (*Little–Rubin* [2002]), ami három változóra vonatkozóan 12 megfigyelést tartalmaz. (A „?” hiányzó adatot jelent.)

Y_1	1	2	3	4	1	2	3	4	?	?	?	?
Y_2	1	2	3	4	?	?	?	?	1	2	3	4
Y_3	?	?	?	?	1	2	3	4	4	3	2	1

Ebben a mintában az elérhető adatpárokat használva a mintából számított korrelációs együtthatók $r_{12} = 1$, $r_{13} = 1$, $r_{23} = -1$. Ezek a becslések nem jók, mert a sokasági korrelációs együtthatóknál $\rho_{12} = \rho_{13} = 1$ -ből az következik, hogy $\rho_{23} = 1$, nem lehet -1 .

Mivel az elérhető adatokat használatával több információra támaszkodunk, azt váránk, hogy ez a megoldás hatékonyabb, mintha csak a teljes adatokat használnánk. *Kim és Curry* [1977] is ezt találták MCAR és gyenge korreláció esetén. Erősebb korreláció esetén viszont a teljes adatok használata bizonyult jobbnak (*Azen–Van Guilder* [1981]).

2.2. Átsúlyozás

Az átsúlyozásos módszerek abból indulnak ki, hogy válaszmegtagadás esetén a válaszmegtagadó elemhez hasonló nem adathiányos esetek (vele azonos kategóriában vagy rétegben szereplő elemek) arányosan több sokasági elemet képviselnek, azaz nagyobb súlyt kell kapniuk. Általában, ha a j -dik alcsoportban (kategóriában) a válaszadók aránya p_j , akkor az itt szereplő elemek $1/p_j$ súlyt kapnak, azaz itt mindegyik elem ennyiszor több sokasági elemet képvisel. Véletlen mintákból való következtetésnél, amikor az elemek kiválasztása nem azonos valószínűséggel történik, gyakran súlyozzák a megfigyelési elemeket a tartalmazási valószínűségük (probability of inclusion, azaz a minták hány százaléka tartalmazza az adott elemet) inverzével (*Hunyadi* [2001]). Legyen például y_i az Y változó értéke az i -dik megfigyelési egységre. Ekkor, ha nincs hiányzó adat, a sokasági átlag Horvitz–Thompson becslőfüggvénye:

$$\hat{Y}_{HT} = \frac{\sum_{i=1}^n \frac{y_i}{\pi_i}}{\sum_{i=1}^n \frac{1}{\pi_i}}, \quad /3/$$

ahol π_i az i -dik egység ismert tartalmazási valószínűsége, a szumma pedig a megkérdezettekre vonatkozik.

Hiányzó adatok esetén az átsúlyozás úgy módosítja a súlyokat, mintha a nemválaszolás is a mintavételi terv része lett volna, ekkor a fenti becslőfüggvény a következőképpen módosul:

$$\hat{Y}_{HTm} = \frac{\sum_{i=1}^n \frac{y_i}{\pi_i \hat{p}_i}}{\sum_{i=1}^n \frac{1}{\pi_i \hat{p}_i}} \quad /4/$$

Itt a szumma nem a megkérdezettekre, hanem a ténylegesen válaszolókra vonatkozik, a \hat{p}_i pedig az i -dik egység becsült válaszadási valószínűsége (általában a válaszadási arány a minta egy alcsoportjában).

A módszer alapelve tehát egyszerű, de többdimenziós feladatoknál már igen bonyolult lehet a kivitelezése. Ráadásul a túlságosan szóródó súlyok nagy korrekciót jelentenek, ami megnöveli a feltételezések szerepét a becslésekben (Hunyadi [2001]). Az átsúlyozás mögött az a feltételezés húzódik meg, hogy az adott rétegen belül a válaszadók a megkérdezettek véletlen almintájának tekinthetők, azaz a rétegen belül az adathiány MCAR-jellegű. Az átsúlyozott mintából sokszor relatíve egyszerű a sokasági paraméterek pontbecsléseit elkészíteni. Az intervallumbecslésekhez szükséges standard hibák számítása már korántsem ilyen egyszerű. A statisztikai programcsomagok lehetővé teszik aszimptotikus standard hibák számítását összetettebb mintavételi tervek esetén, beleértve az átsúlyozást, rétegzést is. Ezek a programok azonban tipikusan fixnek, ismertnek tartják a súlyokat, pedig adathiány esetén a válaszadási aránnyal arányos súlyok maguk is mintavételi ingadozásnak vannak kitéve.

Egyszerű véletlen mintára vannak képletek a hibaszámításhoz, komplexebb esetekhez azonban a minta mesterséges újrahasonosításán alapuló nagy számítógépigényű módszerek (jackknife, bootstrap, kiegyensúlyozott ismétlések) alkalmazására van szükség.

2.3. Imputációalapú eljárások

Az imputáció azt jelenti, hogy a hiányzó adatot utólag mesterségesen pótolják egy ahhoz vélhetően hasonló értékkel. Ezután az így létrejött „teljes” adatbázison elvégezhető a standard statisztikai elemzések. A helyes következtetéshez azonban módosítani kell a standard elemzéseket, valahogyan meg kell különböztetni a valódi és az imputált értékeket, hiszen ez utóbbiak újabb bizonytalansági faktort képeznek. Ezt a bizonytalansági tényezőt építi be a modellbe például a többszörös imputáció (multiple imputation).

Logikai imputációról (data editing) akkor beszélünk, ha a hiányzó értékek más adatokból, vagy korábbi felvételekből logikailag következnek és azokkal pótolják őket. Az emberek neve például nem változik, és a hiányzó életkorra is következtethetünk, ha egy korábbi felmérésnél megadták. A módszer előnye, hogy nem csökkenti az adatokban levő tényleges változékonyságot.

Az átlaggal való pótlás esetében az adott változóban meglévő adatok átlagával (átlag helyett más középérték is használható (módusz, medián)) helyettesítik a hiányzó értékeket. Az átlaggal való imputálás előnye az egyszerűsége, és könnyű alkalmazhatósága. Hátránya viszont, hogy bár teljesen véletlenszerű adathiány esetén várható érték szempontjából nem torzít, az elemek változékonyságát alulbecsli. Ez

javítható, ha a megfigyeléseket homogénebb csoportokra bontjuk és csoportokon belüli részátlagokkal imputálunk, de a standard hibákat és a becslések konfidencia-intervallumát még így is alulbecsüljük. Ez a módszer tulajdonképpen az átsúlyozással azonos eredményt ad.

A *regressziós módszerek* esetében a teljes megfigyeléseken építenek egy regressziót, a hiányzó értéket tartalmazó változót eredményváltozóként, a többi magyarázóváltozóként kezelve. Aztán azokra az esetekre, ahol az eredményváltozó értéke hiányzik, a regresszió segítségével becslést készítenek. A módszer továbbfejlesztéseként a *sztochasztikus regressziós imputálás* esetén egy véletlen változót is adnak a becslésekhez, mert e nélkül a változók közötti kapcsolat a későbbi elemzésekben szorosabbnak mutatkozna, mint amilyen valójában lehet.

A *hot deck imputáció* esetében a hiányzó adatot tartalmazó megfigyeléshez leginkább hasonló hiánymentes esetet megkeresik és ennek Y értékével pótolják a hiányos eset hiányzó Y értékét. A hasonlóság mértékének megítélésére különböző módszerek használhatók. A hot deck módszer előnye a fogalmi egyszerűsége mellett, hogy megőrzi a változók eredeti mérési szintjét (a kategóriás kimenetelű változók kategóriások maradnak, a folytonosak pedig folytonosak). A módszer hátránya, hogy nehéz az esetek hasonlóságát definiálni és az elemzőnek esetleg saját programot kell készítenie a donor egységek kiválasztásához. Ezenkívül a standard hibák számítása is nehézségekbe ütközhet (Roth–Switzer [1995]). A nehézségek ellenére a hot deck imputáció igen népszerű technika, számos hivatalos statisztikai felmérésben is ezt a módszert alkalmazták. (Például: Statistics Canada (Rubin [1987]).) Vannak modellek, amelyek több hasonló esetet keresnek és azokból véletlenszerűen választják ki a donor megfigyelést, vagy ha az megfelelő, az átlagukat számítják az imputációhoz. A hot deck (belső) módszereken sokszor tágabb értelemben az olyan adatpótlást értik, amely csak az adott mintát használja az imputációhoz, cold deck (külső) módszerek esetén pedig más, külső forrásokat (az adott mintához képest külső, például múltbeli hasonló felmérések adatai) is felhasználnak.

A *közelítő bayesi bootstrap* (Approximate Bayesian Bootstrap – ABB) módszer logisztikus regressziót alkalmaz, hogy az Y függő változóban a válaszolás/nemválaszolás valószínűségét becsülje az X_i változók segítségével. (Ilyen logisztikus regressziós módszert alkalmaz György [2004] a munkaerő-felvételben szereplő nemválaszolás kezelésére.) A megfigyelési egységek az így kapott hiányzás hajlamossági score-ok alapján képzett kvantilisokba csoportosíthatók. A csoportokon belül a nem hiányos esetekből visszatevéses mintavétel segítségével lehet imputálni a hiányzó értékeket. Az eljárás minden hiányzó adatot tartalmazó változóra megismétlődik. A módszer a hot deck imputáció egy formája, ahol a hasonlóságot a hiányzás hajlamossági score-ok határozzák meg.

Léteznek ún. *kompozit módszerek* (composite methods) is, amelyek különböző módszerek alapelemeit ötvözik. Például a hot deck és a regressziós imputáció keve-

réke, amely először regresszióval számítja a becsült átlagokat, majd ezekhez hozzáadja egy véletlenszerűen kiválasztott empirikus reziduum értékét.

A nemválaszolás miatti bizonytalanság pótlólagos varianciaforrást jelent, amit valahogyan be kell építeni a becslésekbe. Ez megoldható például a minták másodlagos hasznosításán alapuló számítógép intenzív módszerek alkalmazásával, amelyekkel bonyolult mintavételi terv és imputációs technika esetén is becsülhető a becslőfüggvények varianciája. Több imputált adatbázis létrehozásával és azok eredményeinek összesítésével szintén beépíthető az adathiány okozta pótlólagos bizonytalanság a rendszerbe.

A többszörös imputáció (Multiple Imputation – MI) esetén minden hiányzó elem helyére több lehetséges értéket imputálnak, ezáltal több (általában 3-10) „teljes” adatbázist készítenek az eredeti hiányos adatbázisból. Az elemző mindegyik adatbázison elvégzi a megfelelő statisztikai módszerekkel a kívánt elemzéseket, a kapott eredményeket összegyűjti és kombinálja egyetlen elemzésbe. Ez utóbbi sokszor nem egyszerű feladat. A többszörös imputáció egy lépéssel tovább megy azzal, hogy bevezeti a statisztikai bizonytalanságot a modellbe, azért, hogy egy teljes adatbázisban meglévő változékonyságot közelítse az imputációval teljessé tett adatbázis is.

A többszörös imputációt először *Rubin* [1987] javasolta a hiányzó adatok kezelésére.

Furcsának tűnhet, hogy viszonylag kevés (3–10) imputációval is érzékeltetni lehet a pótlások bizonytalanságát. *Rubin* megmutatta, hogy m imputáción alapuló becslés relatív hatékonysága végtelen számú imputáció hatékonyságához képest nagyjából

$\left(1 + \frac{\gamma}{m}\right)^{-1}$, ahol γ a hiányzó információk aránya (számítását lásd később). Az m

és γ különböző értékei mellett elérhető hatékonyságokat mutatja az alábbi táblázat:

Többszörös imputációval elérhető relatív hatékonyság
(százalék)

m	γ				
	0,1	0,3	0,5	0,7	0,9
3	97	91	86	81	77
5	98	94	91	88	85
10	99	97	95	93	92
20	100	99	98	97	96

Ha a hiányzó információk aránya nem túl magas, akkor igen kevés javulást eredményez néhánynál több imputált adatbázis készítése és elemzése. Az m darab imputált adatbázison elvégzett elemzések eredményeinek összegzésére *Rubin* azt a

módszert ajánlotta, hogy minden elemzésből mentsük el a becsült paraméterek és a standard hibák értékét. Legyen $\hat{\theta}_j$ a becsülni kívánt paraméter értéke (például egy regressziós együttható) a j -edik adathalmazból ($j=1,2,\dots,m$). U_j pedig legyen a $\hat{\theta}_j$ varianciája. Az összesítés utáni becslés az egyedi becslések átlaga lesz:

$$\bar{\theta} = \frac{1}{m} \sum_{j=1}^m \hat{\theta}_j \quad /5/$$

Ezen becslés standard hibájához először az átlagos imputáción belüli varianciát:

$$\bar{U} = \frac{1}{m} \sum_{j=1}^m U_j \quad /6/$$

és az imputációk közötti varianciát kell kiszámolni:

$$B = \frac{1}{m-1} \sum_{j=1}^m (\hat{\theta}_j - \bar{\theta})^2. \quad /7/$$

A teljes variancia:

$$T = \bar{U} + \left(1 + \frac{1}{m}\right) B. \quad /8/$$

Ahol az $\left(1 + \frac{1}{m}\right)$ a véges m miatti korrekciós tényező.

Az együttes standard hiba pedig \sqrt{T} lesz.

A $\gamma = \frac{(1+m^{-1})B}{T}$ a nemválaszolás miatt a θ -ról hiányzó információk becsült aránya.

Nagy minták esetén a θ -ra vonatkozó szignifikancia tesztelése a $t = \frac{\theta - \bar{\theta}}{\sqrt{T}}$ próba-függvénnyel történhet, ami a nullhipotézis alatt Student-féle t -eloszlást követ a következő szabadságfokkal:

$$v = (m-1) \left(1 + \frac{1}{m+1} \frac{\bar{U}}{B}\right)^2 \quad /9/$$

ami a Satterthwaite-közelítésen alapul (*Rubin–Schenker* [1986] és *Rubin* [1987]).

A szabadságfok javított értéke kis mintákra:

$$v' = (v^{-1} + \hat{v}_{megfigy}^{-1})^{-1}, \quad /10/$$

ahol

$$\hat{v}_{megfigy} = (1 - \gamma) \left(\frac{v_{telj} + 1}{v_{telj} + 3} \right) v_{telj}, \quad /11/$$

és v_{telj} az adathiányt nem tartalmazó adatbázis esetén alkalmazandó szabadságfok. (*Barnard–Rubin* [1999]).²

Intervallumbecslés a paraméterre szintén ezek felhasználásával készülhet. További módszereket ismertet az eredmények összesítésére többszörös imputáció esetén *Schafer* ([1997], 4. fejezet).

A többszörös imputáció különböző módszerekkel történhet attól függően, hogy milyen jellegzetességekkel bír az adathiány. A többszörös imputációt besorolhatnánk a modellalapú eljárások közé is, mert legtöbbször bayesi eljárásról alapul: szükség van egy parametrikus modellre a teljes adatokra vonatkozóan és prior eloszlásra az ismeretlen modell paraméterekre (ez esetlegesen lehet neminformatív), aztán a hiányzó adatokra készít m független szimulációt a hiányzó adatok feltételes eloszlását használva (Bayes-tétel). Bonyolultabb parametrikus modellek esetén speciális számítási technikákra is szükség lehet, ezek közül leggyakrabban a Markov-lánc Monte-Carlo³ (Markov chain Monte Carlo – MCMC) szimulációt használják. *Rubin* [2003] egy olyan MCMC-szimulációt és beágyazott többszörös imputációt alkalmazó modellt ír le, amelyet három változó esetén a következő módon lehet illusztrálni.

Legyen a három változónk X , Y és Z . Kezdjük azzal, hogy valahogyan kitöltjük az Y és Z hiányzó értékeit (ezek az induló értékek), majd a megfigyelt X -ekkel építünk egy $X|Y, Z$ modellt és e modell segítségével imputáljuk a hiányzó X -eket. Ezek után dobjuk ki az imputált (induló) Y értékeket és illesszünk egy $Y|X, Z$ modellt a megfigyelt Y -okra, majd ezzel a modellel imputáljuk a hiányzó Y -okat. Aztán dobjuk ki az imputált Z értékeket és illesszünk egy $Z|X, Y$ modellt a megfigyelt Z -kre, majd ezzel a modellel imputáljuk a hiányzó Z -ket. Az iteratív eljárás mindaddig folytatja a fenti lépések ismétlését, míg a kapott paraméterek nem konvergálnak.

² A képletek pontos elméleti háttere megtalálható a hivatkozott művekben.

³ A Markov-lánc véletlen változók sorozata, amelyben minden egyes elem eloszlása az előző értékétől függ. A módszert eredetileg a fizikában használták egymással kölcsönhatásba lépő molekulák egyensúlyi eloszlásának feltárására. A statisztikai alkalmazások során többdimenziós, más módszerekkel megfoghatatlan eloszlások generálására használják.

A többszörös imputáció előnye, hogy könnyen érthető és elég robusztus a változók normalitási feltételének sérülése esetén is. Még például a bináris vagy az ordinális skálán mérő kategóriás változók esetén is gyakran elfogadható a normalitási feltétel melletti imputáció, majd a kapott folytonos imputált érték kerekíthető a legközelebbi kategóriára. Az erőteljes aszimmetriával rendelkező eloszlások közel normálissá transzformálhatók (például logaritmizálással), majd imputáció után visszatranszformálhatók az eredeti skálára. Hátránya viszont, hogy időigényes a három-tíz adatbázis imputálása, majd külön-külön az elemzések elvégzése, végül ezek összegzése. Ráadásul az összegzés módszertana még nincs minden statisztikai modellre kidolgozva. A többszörös imputációt több statisztikai szoftverbe is beépítették (például: a SAS enterprise Miner-hez írt Intelligent Multiple Imputation Software System – IMISS) ezek használatával az eljárás időigénye csökkent és sok kutató számára vonzó megoldássá vált.

2.4. Modellalapú eljárások

A modellalapú eljárások egy modellt definiálnak a megfigyelt adatokra és a becsléseket a modell melletti posterior valószínűségekre, vagy likelihoodra alapozzák. A megközelítés előnye a rugalmasság, a modellnél alkalmazott feltételezések explicit volta és az adathiányt is beépítő varianciabecslések elérhetősége.

Ilyen modellalapú becslés a maximum likelihood (ML) becslés, ami kiváló nagymintás tulajdonságokkal rendelkezik (konzisztens, aszimptotikusan hatásos, határeloszlása normális) (Hunyadi–Vita [2002]). A hiányzó adatok mintázata azonban nem mindig teszi lehetővé az ML-becslések explicit számítását.

Tegyük fel, hogy van egy modellünk az Y -ra, melynek eloszlását az $f(Y|\theta)$ sűrűségfüggvénnyel írhatjuk le, ahol θ ismeretlen paraméter. Legyen $Y = (Y_{\text{megfigyelt}}, Y_{\text{hiányzó}})$, ekkor $f(Y|\theta) = f(Y_{\text{megfigyelt}}, Y_{\text{hiányzó}} | \theta)$ az $Y_{\text{megfigyelt}}$ és az $Y_{\text{hiányzó}}$ együttes eloszlását leíró sűrűségfüggvény, az $Y_{\text{megfigyelt}}$ peremeloszlása pedig :

$$f(Y_{\text{megfigyelt}}|\theta) = \int f(Y_{\text{megfigyelt}}, Y_{\text{hiányzó}} | \theta) dY_{\text{hiányzó}}$$

Ekkor MAR-adathiány esetén a likelihood:

$$L(\theta | Y_{\text{megfigyelt}}) = \int f(Y_{\text{megfigyelt}}, Y_{\text{hiányzó}} | \theta) dY_{\text{hiányzó}}$$

Ekkor a ML-becslés a következő egyenlet megoldásával kapható:

$$D_{\ell}(\theta | Y_{\text{megfigyelt}}) = \frac{\partial \ln L(\theta | Y_{\text{megfigyelt}})}{\partial \theta} = 0$$

Ha ennek az egyenletnek nincs zárt alakú megoldása, akkor iteratív módszerek alkalmazására van szükség. Ilyen iteratív módszer például a Newton–Raphson-algoritmus. Egy alternatív módszer a hiányzó adatokkal való becslések készítéséhez a várakozás maximalizáció (expectation maximization – EM), ami nem igényli a második deriváltak számítását, így nincs szükség olyan komplex programozási megoldásra, mint a Newton–Raphson-algoritmust alkalmazó módszerek esetén.

A következőkben ezt a módszert mutatjuk be, mert a gyakorlatban nagyon elterjedt az alkalmazása.

A várakozás maximalizáció egy általános módszer maximum likelihood becslésre MAR-típusú adathiány esetén. A módszer egy iteratív eljárás, amely két lépésből áll. Először, a várakozási lépésben (E) kiszámítják a teljes adatokat tartalmazó állományra a loglikelihood várható értékét, azután a maximalizáló lépésben (M) a kapott várható értékeket behelyettesítik a hiányzó értékek helyére és maximalizálják a likelihood függvényt, mintha nem lett volna hiányzó adat. Így új paraméterbecsléseket kapnak. Ez az iteratív eljárás mindaddig folytatja a fenti két lépés ismétlését, míg a kapott paraméterek nem konvergálnak. Konvergenciáról akkor beszélhetünk, ha a paraméterbecslések változása lépésről lépésre egyre kisebb lesz mígnem teljesen elhanyagolhatóvá válik. A konvergenciához annál több iteráció szükséges, minél több a hiányzó adat.

Nézzük meg egy egyszerű példán, hogyan működik az EM-módszer. A becslés elvégzéséhez valójában nincs szükség az EM-algoritmusra, csak a szemléltetés kedvéért választottuk.

Tegyük fel, hogy négyszer egymás után feldobunk egy pénzérmét, aminek az eredménye: (fej, fej, írás, ?), ahol a ? azt jelenti, hogy a negyedik dobás eredményét valamilyen oknál fogva nem ismerjük. Legyen a becsülni kívánt sokasági paraméter a „fej-dobás” valószínűsége, π . A teljes Y adatállományt felbontjuk megfigyelt és hiányzó részre: $Y = (Y_{\text{megfigyelt}}, Y_{\text{hiányzó}})$, a megfigyelt adatok valószínűségét a következő módon kapjuk:

$$\begin{aligned} P(Y_{\text{megfigyelt}}|\pi) &= \sum_{Y_{\text{hiányzó}}} P(Y|\pi) = \\ &= P((F,F,\acute{I},\acute{I})|\pi) + P((F,F,\acute{I},F)|\pi) = \pi^2(1-\pi)^2 + \pi^3(1-\pi) = \pi^2(1-\pi) \end{aligned}$$

A megfigyelt adatok valószínűsége tehát ugyanaz, mintha a negyedik dobást egyáltalán nem vennénk figyelembe. Ekkor a π maximum likelihood becslése:

$$L(\pi|Y_{\text{megfigyelt}}) = P(Y_{\text{megfigyelt}}|\pi) = \pi^2(1-\pi)$$

$$D_L(\pi|Y_{\text{megfigyelt}}) = \frac{\partial L(\pi|Y_{\text{megfigyelt}})}{\partial \pi} = 2\pi(1-\pi) - \pi^2 = 2\pi - 3\pi^2 = 0 \rightarrow \hat{\pi}_{ML} = \frac{2}{3}$$

A szemléltetés kedvéért nézzük, hogyan kaptuk volna meg ezt az eredményt az EM-módszer segítségével! Az E várakozási lépésben felírjuk a teljes adatok loglikelihoodjának várható értékét a jelenlegi $\pi^{(t)}$ becslés mellett.

$$Q(\pi | \pi^{(t)}) = \pi^{(t)}(3\ln\pi + \ln(1-\pi)) + (1-\pi^{(t)})(2\ln\pi + 2\ln(1-\pi))$$

Az M maximalizálási lépésben keressük Q maximumát π szerint, hogy megkapjuk $\pi^{(t+1)}$ -et.

$$D_Q(\pi | \pi^{(t)}) = \frac{\partial Q(\pi | \pi^{(t)})}{\partial \pi} = \pi^{(t)}\left(\frac{3}{\pi} - \frac{1}{1-\pi}\right) + (1-\pi^{(t)})\cdot\left(\frac{2}{\pi} - \frac{2}{1-\pi}\right) = 0$$

Ebben az egyszerű esetben zárt formát kapunk az iterációra: $\pi^{(t+1)} = 0,5 + 0,25\pi^{(t)}$

Ha a kiinduló becslésünk mondjuk $\pi^{(0)} = 0,25$, akkor az iterációk sorozata: 0,2500; 0,5625; 0,6406; 0,6602; 0,6650; 0,6663; ..., ami konvergál a 2/3-hoz.

Az EM-megközelítés előnye, hogy jól ismert statisztikai tulajdonságai vannak és általában jobban működik, mint az egyszerűbb listwise és pairwise adattörlések, az átlaggal való helyettesítés, vagy a regressziós imputálás (*Little* [1979], *Donner–Rosner* [1982], *Lee–Chiu* [1990]). Monte-Carlo-szimulációk is hasonló eredményeket mutattak (*Malhotra* [1987], *Graham–Donaldson* [1993]). Ugyanakkor ez az előny sokszor igen kicsi lehet (*Donner–Rosner* [1982]). A módszer hátránya annak viszonylagos bonyolultsága, ami miatt inkább csak statisztikusok számára vonzó megoldás. A legfontosabb gyengéje a módszernek, hogy a becsült adathoz nem ad bizonytalansági komponenst. A gyakorlatban ez azt jelenti, hogy míg a paraméterbecslések torzítatlanok lesznek, addig a standard hibák és a kapcsolódó tesztek nem megbízhatók. Ez a hiányosság arra készítette a statisztikusokat, hogy újabb likelihood alapú módszereket fejlesszenek ki. Ilyenek a teljes információs maximum likelihood módszer vagy a fent már tárgyalt többszörös imputáció alkalmazása.

(A teljes információs maximum likelihood (Full Information Maximum Likelihood – FIML vagy Raw Maximum Likelihood) minden elérhető adatot használ, hogy maximum likelihood alapú becsléseket készítsen. A módszert részletesen ismerteti például *Wothke* [1998].)

A maximum likelihood módszer MAR-típusú adathiányt feltételez, de a listwise és pairwise törlésekhez képest még nem véletlenszerű adathiány esetében is jobb eredményeket ad (*Wothke* [1998]).

A korábban ismertett eljárások alkalmazásának szükséges feltétele a véletlenszerű adathiány (MAR). Vannak azonban olyan körülmények, amelyek esetén ez a feltételezés nem tartható, mert az adathiány kapcsolatban van a hiányt tartalmazó változóval. Ekkor az adathiány jellegét figyelembe vevő, a nem véletlenszerű adathiány kezelésére szolgáló modellek alkalmazására van szükség.

A NMAR-adathiánnyal foglalkozó kutatások alapvetően eltérő megközelítésük alapján két csoportra bonthatók: *szelekciós modellek* és *mintázatkeverék- (pattern-mixture) modellek*. Ezek a modellek az együttes valószínűséget eltérő módon bontják fel. A *szelekciós modellek* a $P(y_{hiányzó}, y_{megfigyelt}) = P(y_{hiányzó} | y_{megfigyelt}) P(y_{megfigyelt})$ felbontást használják. A szelekciós modellek feltételezik, hogy az adathiányt tartalmazó változó akkor és csak akkor figyelhető meg, ha egy másik változó (ami nem megfigyelhető) átlép egy küszöbértéket. Ilyen módszert alkalmazott Heckman [1976] kétlépcsős probit modelljében. A szelekciós modellek esetén a likelihood szokatlan eloszlású lehet, mert a paraméterek becsléséhez sokszor kevés információ áll rendelkezésre (Schafer–Graham [2002]).

A megoldás alternatívájaként alkalmazhatók a *mintázatkeverék-modellek*, amelyek a $P(y_{hiányzó}, y_{megfigyelt}) = P(y_{megfigyelt} | y_{hiányzó}) P(y_{hiányzó})$ felbontást alkalmazzák.

A mintázatkeverék-modellekkel foglalkozó tanulmányok: Hedeker–Gibbons [1997], Little–Schenker [1994], Little [1993], és Glynn–Laird–Rubin [1986]. Ezek a modellek kategorizálják a hiányzó értékek különböző mintázatait egy magyarázó változóba és ezt a magyarázó változót beépítik az adott statisztikai modellbe. Ezek után meghatározható, hogy az adathiány jellegzetességének van-e prediktív ereje akár önállóan (közvetlen hatás), akár más változókkal együttesen (interakciós hatás). A módszer előnye, hogy nem feltételezi a véletlenszerű adathiányt és részben használható hozzá statisztikai szoftverek, például a SAS MIXED proc. (például Hedeker–Gibbons [1997]), hátránya viszont, hogy az elemzőnek magának kell bizonyos lépéseket leprogramozni. Ha a megfigyelések számához képest sok változó esetén van relatíve sokféle eredetű adathiány, akkor a módszer elegendő adat hiányában nem működik.

3. Összegzés

A hiányzó adatok kezelésére nem létezik tehát egyetemesen legjobb megoldás. Pontosabban a legjobb gyógymód itt is a megelőzés. Ez sajnos nem mindig lehetséges, így ha már van adathiány, és az nem teljesen véletlenszerű, akkor valamilyen módon kezelni kell.

Összességében elmondható, hogy az általánosan használt egyszerű adathiány kezelési eljárásoknál (listwise és pairwise törlés, átlag imputálás) a hot deck, a maximum likelihood alapú és a többszörös imputációs eljárások a legtöbb esetben jobban teljesítenek. Mivel egyre szélesebb körben elérhető és könnyen használható szoftverek is tartalmazzák ezeket az eljárásokat, így az elméleti szerepükön túl az alkalmazásuk is egyre gyakoribb. Ezen módszerek mindegyike feltételezi a véletlenszerű

adathiányt, vannak azonban újabb statisztikai modellek a nem véletlenszerű adathiány kezelésére is. Ezekhez is használhatók (részben) az ismert statisztikai programcsomagok.

Az eljárások közötti választásban fontos szerepe van annak, hogy a cél *parameterbecslések és tesztstatisztikák készítése*, vagy *konkrét megfigyelések hiányzó adatainak becslése*. Az első esetben az adatbázis felhasználója kezeli a hiányzó adatokat és választhatja a saját elemzéséhez leginkább megfelelő módszert. A második esetben, ha például statisztikai hivatalok, kormányzati szervek nyilvánosságnak szánt adatbázisairól van szó, vagy olyan vállalati adatbázisokról, amelyeket sokféle belső kutatáshoz használnak, akkor olyan megoldást kell választani, ami nem igényel túl komplex bánásmódot a végső elemzések elvégzésekor. Ekkor például nem nagyon alkalmazható a többszörös imputáció. Fontos, hogy a választott imputációs eljárás kompatibilis legyen az imputált adatbázison később elvégzendő elemzésekkel. Az imputációs modellel szembeni elvárás, hogy megőrizze a későbbi vizsgálat tárgyát képező változók közötti kapcsolatokat. Ha például az Y változót egy olyan modellel imputálták, amelyik csak az X_1 változót tartalmazta, majd imputáció után a kutató egy lineáris regressziós modellt illeszt Y -ra X_1 és X_2 változók felhasználásával, akkor az X_2 együtthatója torzított lesz 0 felé, a helytelen imputáció következtében. Hasonló okokból, panel felvételeknél, például a keresztmetszeti kapcsolatokon kívül az adott változó korábbi hullámbeli tényleges vagy imputált értékét is figyelembe kell venni. Az imputált adatbázisokhoz mellékelni kell az imputáló által alkalmazott modellt, mert így az elemző láthatja, hogy milyen változókat vontak be a modellbe és mely változók közötti kapcsolatokat tekintettek impliciten 0-nak.

Az imputációt sokan egyfajta *statisztikai alkímiának* tartják, amelyben a semmiből valahogyan új információ keletkezik. Ez a felvetés helytálló lehet az olyan imputációs eljárásokkal kapcsolatban, amelyek az imputált értékeket ugyanúgy kezelik, mint a ténylegesen megfigyelteket. Ha viszont pontosan közlik az alkalmazott módszert és a hiányzó adatok bizonytalansága is megjelenik, akkor a hiányzó adatok megfelelő kezelésével eltüntethető, vagy legalábbis csökkenthető a nem teljesen véletlen adathiányból eredő torzítás.

Irodalom

- AZEN, S – VAN GUILDER, M. [1981]: Conclusions regarding algorithms for handling incomplete data. *Proceedings of the Statistical Computing Section, American Statistical Association*. 53–56. old.
- BARNARD, J. – RUBIN, D. B. [1999]: Small-sample degrees of freedom with multiple imputation. *Biometrika*. 86. évf. 4. sz. 949–955. old.
- CSEREHÁTI Z. [2004]: Az outlierek meghatározása és kezelése a gazdaságstatisztikai felvételekben. *Statisztikai Szemle*. 82. évf. 8. sz. 728–746. old.

- DEMPSTER, A. P. – LAIRD, N. M. – RUBIN, D. B. [1977]: Maximum likelihood estimation from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society B*. 39. évf. 1. sz. 1–38. old.
- DONNER, A. – ROSNER, B. [1982]: Missing value problems in multiple linear regression with two independent variables. *Communication in Statistics*. 11. évf. 2. sz. 127–140. old.
- GLYNN, R. – LAIRD, N. M. – RUBIN, D. B. [1986]: Selection modeling versus mixture modeling with nonignorable nonresponse. In: *Wainer, H. (szerk.): Drawing Inferences from Self-Selected Samples*. Springer-Verlag, New York.
- GRAHAM, J. W. – DONALDSON, S. I. [1993]: Evaluating interventions with differential attrition: the importance of nonresponse and use of followup data. *Journal of Applied Psychology*. 78. évf. 1. sz. 119–128. old.
- GYÖRGY E. [2004]: A nemválaszolás elemzése a munkaerő felvételnél. *Statisztikai Szemle*. 82. évf. 8. sz. 747–772. old.
- HECKMAN, J. J. [1979]: Sample selection bias as a specification error. *Econometrica*. 47. évf. 1. sz. 153–161. old.
- HEDEKER, D. – GIBBONS, R. D. [1997]: Application of random-effects pattern-mixture models for missing data in longitudinal studies. *Psychological Methods*. 2. évf. 1. sz. 64–78. old.
- HUNYADI L. [2001]: A mintavétel alapjai. *Egyetemi Jegyzet SZÁMALK*. Budapest.
- HUNYADI L. – VITA L. [2002]: Statisztika közgazdászoknak. *Központi Statisztikai Hivatal*. Budapest.
- KIM, J. O. – CURRY, J. [1977]: The treatment of missing data in multivariate analysis. *Sociological Methode Recherche*. 6. évf. 2. sz. 215–240. old.
- LEE, S. Y. – CHIU, Y. M. [1990]: Analysis of multivariate polychoric correlation models with incomplete data. *British Journal of Mathematical and Statistical Psychology*. 43. évf. 1. sz. 145–154. old.
- LITTLE, R. J. A. [1979]: Maximum likelihood inference for multiple regression with missing values: a simulation study. *Journal of the Royal Statistical Society*. 41. évf. 1. sz. 76–87. old.
- LITTLE, R. J. A. – RUBIN, D. B. [1987]: *Statistical analysis with missing data*. John Wiley & Sons. New York.
- LITTLE, R. J. A. – RUBIN, D. B. [2002]: *Statistical analysis with missing data*. 2. szerk. John Wiley & Sons. New York.
- LITTLE, R. J. A. [1993]: Pattern-mixture Models for Multivariate Incomplete Data. *Journal of the American Statistical Association*. 88. évf. 421. sz. 125–134. old.
- LITTLE, R. J. A. – SCHENKER, N. [1994]: Missing Data. In: *Arminger, G. – Clogg, C. C. – Sobel, M. E. (szerk.): Handbook for Statistical Modeling in the Social and Behavioral Sciences*. Plenum. New York. 39–75. old.
- MÁDER M. P. [2005]: Imputálási eljárások hatékonysága. *Statisztikai Szemle*. 83. évf. 7. sz. 628–644. old.
- MALHOTRA, N. K. [1987]: Analyzing marketing research data with incomplete information on the dependent variable. *Journal of Marketing Research*. 24. évf. 1. sz. 74–84. old.
- ROTH, P. L. – SWITZER, F. S. [1995]: A Monte Carlo analysis of missing data techniques in a HRM setting. *Journal of Management*. 21. évf. 5. sz. 1003–1023. old.
- RUBIN, D. B. [1987]: *Multiple imputation for nonresponse in surveys*. John Wiley & Sons. New York.

- RUBIN, D. B. [2003]: Nested multiple imputation of NMES via partially incompatible MCMC. *Statistica Neerlandica*. 57. évf. 1. sz. 3–18. old.
- RUDAS T. [1998]: *Hogyan olvassunk közvélemény-kutatásokat?* Új Mandátum Könyvkiadó. Budapest.
- SCHLAFER, J. L. [1997]: *Analysis of incomplete multivariate data*. Chapman & Hall. London.
- SCHAFFER, J. L. – GRAHAM, J. W. [2002]: Missing data: our view of the state of the art. *Psychological Methods*. 7. évf. 2. sz. 147–177. old.
- STATISTICAL SOLUTIONS, Inc. [1998]: SOLAS for missing data analysis. Version 1. Cork, Ireland: Statistical Solutions.
- WOTHKE, W. [1998]: Longitudinal and multi-group modeling with missing data. *Mahwah, NJ: Lawrence Erlbaum Associates*.

Summary

Missing data cause several problems. Inconsistency can be experienced among the results of analyses done on the same database by different researchers, if they handled missing data in a different way. There is a need of handling missing data, because the populational parameter estimations may be biased, unless the missing of data is not completely at random.

The aim of handling missing data is exactly to make this bias disappear. Different methods reach this aim in the function of how correctly the features of missing data can be identified and constructed. In this article we look shortly over the types of missing data and the most often recommended methods used to handle them, highlighting their main advantages and disadvantages.

There is no universally best solution to handle missing data. But we can say, that methods, based on hot deck, maximum likelihood and multiple imputation methods usually perform better, than generally used methods, handling simple missing of data (listwise and pairwise deletion, mean imputation). As widely accessible and easily practicable software include these methods, their application over their theoretical role is more and more common. All these methods presume missing data at random, there are however newer statistical methods to handle data, not missing at random as well.

Inna Šteinbuka,
az Eurostat D Igazgatóságának
igazgatója
E-mail: inna.steinbuka@ec.europa.eu

Pascal Wolff
az Eurostat főmunkatársa,
E-mail: pascal.wolff@ec.europa.eu,

Mutatók és jobb politikaalkotás: a fenntartható fejlődés kérdésköre*

A fenntartható fejlődési stratégiák (Sustainable development strategies – SDS) alapja az Agenda 21, mely az 1992. évi Riói Csúcstalálkozón elfogadott, a fenntartható fejlődés elérésére vonatkozó program alapdokumentuma. Az Agenda 21 felszólított minden országot, hogy az érdekelt felek széles körének bevonásával fejlesszék ki a stratégiát. A 2001-es kezdeti kötelezettségvállalást követően 2006 júniusában az Európai Tanács elfogadta a megújított fenntartható fejlődési stratégiát¹, mely egy egységes dokumentumban írja le az Európai Unió hosszú távú jövőképét. Ahogy azt eredetileg az Agenda 21 40. fejezetében rögzítették, az EU stratégiája elismeri a statisztikai mutatók szerepét, és hogy fontos a fenntartható fejlődés előrehaladásának mutatószámok alapján történő nyomon követése.

A jelen dolgozat leírja, hogy a hivatalos statisztika milyen szerepet játszhat egy olyan összetett törekvésben, mint a fenntartható fejlődés. A kiindulási pontot – az EU-stratégiát és a mutatók elismert szerepét – az első rész tartalmazza. A második rész áttekintést ad a mutatók alkalmazási szakaszáról, külön hangsúlyt fektetve az

* A jelen dolgozatban kifejtett nézetek a szerzők véleménye, és nem feltétlenül tükrözik az Európai Bizottság nézeteit.

A cikk a 2007. augusztus 22. és 29. között Lisszabonban tartott, Nemzetközi Statisztikai Intézet 56. konferenciáján szervezett „Mire van szükségük a felhasználóknak? Mutatószámokra vagy statisztikára” c. különleges ülésre benyújtott prezentáció szerkesztett változata. A fordítást Szabóné Radnóti Zsuzsanna, a KSH vezető főtanácsosa készítette.

¹ Európai Tanács, 2006. június 26-i 10917/06 számú dokumentuma: „Az EU fenntartható fejlődési stratégiájának felülvizsgálata – megújított stratégia”.

elemző munkára, míg a harmadik rész néhány, a gyakorlat során nyert tapasztalatot mutat be.

A politika és a statisztikai folyamatok összekapcsolása

A 2006-os megújított stratégia fontos előrelépést jelentett a fenntartható fejlődési politikák megfogalmazása és az irányukba való elkötelezettség felé az Európai Unióban. A monitoring tekintetében elismerte a mutatók szerepét, ezzel erősítve mind a fontosságukat, mind a statisztikai közösség szerepét a stratégia monitoringjában.

Az EU fenntartható fejlődési stratégiája. Bár az Európai Közösséget létrehozó szerződés a fenntartható fejlődést egy átfogó, minden EU-politikát alátámasztó fogalomként határozza meg, az Európai Tanács által 2001 júniusában elfogadott fenntartható fejlődési EU-stratégia előtt nem volt határozott állásfoglalás annak megvalósításáról. A 2001-es SDS nem volt egységes dokumentum, következésképpen hiányzott belőle a világosság és átláthatóság. Hozzá kell tenni, hogy a stratégia nem tartalmazott világos szabályokat az irányítás és a tulajdonjog vonatkozásában sem.

Az Európai Tanács viszont 2006-ban elfogadott egy egységes, összefüggő stratégiát arról, hogy az EU hogyan tud hatékonyabban megfelelni régóta fennálló, a fenntartható fejlődés kihívásaival kapcsolatos kötelezettségeinek (4. bekezdés). Az EU-SDS-ben megfogalmazott célok és célkitűzések megvalósítása erőfeszítéseket tesz szükségessé mind európai szinten, mind az EU-tagországok részéről. Az EU-SDS egyik vezérlőelve így a különböző szintű politikaalkotási szintek (EU, nemzeti, helyi) közötti koherencia megteremtése a stratégia megvalósítása érdekében. A megújított stratégia másik jelentős eredménye a Lisszaboni Stratégiával való együttműködés tisztázása, mely a fenntartható fejlődés átfogó célkitűzéseire elsősorban a gazdasági fejlődésre és munkahelyteremtésre koncentrálni járul hozzá.

A megújított SDS bevezet egy irányítási ciklust is (33-45. bekezdés), mely tisztázza az ütemezést és a legfontosabb érdekelt felek szerepét. Nevezetesen, az Európai Bizottság két évente fejlődési jelentést készít az SDS megvalósulásáról az EU-ban és a tagországokban. A jelentés a fenntartható fejlődés mutatókészletére (Sustainable development indicators – SDI) és az Eurostat azonos periodicitást követő monitoringjelentésére támaszkodik. Ha a stratégia meg is határozza a kidolgozási folyamatot és mutatók használatát a jelentés készítésében, magát a mutatólistát nem írja elő, bár tervezték, hogy 2007. végére az Európai Tanács elfogad egy korlátozott számú mutatókészletet. Ezzel a megoldással a legtöbb érdekelt fél egyetértett, mivel így elkerülhető a mutatók befagyása, és lehetséges a rugalmasabb javításuk és fejlesztésük az idők során.

A mutatók iránti igény a stratégiában. Az Agenda 21 40. fejezete elismerte a döntéshozatalhoz szükséges információk iránti igényt, és felhívta az országokat és a

nemzetközi közösséget, hogy fejlesszék ki a fenntartható fejlődés mutatóit. Ennek szellemében a stratégia megköveteli olyan statisztikai mutatók kidolgozását, melyek a fenntartható fejlődést összességében lefedik, és lehetővé teszik a haladás megfelelő értékelését.

Az SDS-teszt néhány általános ajánlást a mutatók által lefedett területekről (34. bekezdés) és a mutatókészlet folyamatos javításának szükségességéről. A mutatók kidolgozása az Európai Bizottság feladata a tagországokkal együttműködve egy szakértői SDI-munkacsoport segítségével. Ez a munkacsoport segíti a Bizottságot a legrelevánsabb mutatók kiválasztásában. A stratégia számít a meghatározott mutatókészlet politikai jóváhagyására és elfogadására is (36. bekezdés). Végül, de nem utolsósorban a stratégia utal a mutatók felhasználására mind az Eurostat monitoringjelentésében, mind a Bizottság fejlődési jelentésében.

A legfontosabb érintettek jóváhagyása számos okból kifolyólag kívánatos és fontos tényező a sikeres kimenetel szempontjából. A jóváhagyás és az érintettekkel való konzultáció hiánya azt jelentheti, hogy sokéves munkával olyasmit állítanak elő, amit esetleg elismernek, de soha nem használnak.

Az európai statisztikai rendszer reagálása

A 2001-es stratégia elfogadása ahhoz vezetett, hogy az Eurostat felismerte, hogy szükséges egy kifejezetten az EU-ra szabott mutatókészlet kidolgozása. 2001 szeptemberében létrehoztak egy munkacsoportot, hogy kidolgozza az Európai Statisztikai Rendszer közös megközelítését a fenntartható fejlődés mérésének vonatkozásában. Ez a munkacsoport 16 európai ország szakértőjéből és a Bizottság 12 szakpolitikai főigazgatóságából állt. Néhány elkerülhetetlen kompromisszum ellenére a munkacsoport eredményeit 2005-ben a Bizottság összességében jóváhagyta *José Almunia* úr, az EU Gazdasági és Pénzügyek biztosa közleménye által.²

2006 novemberében az Eurostat az SDI-munkacsoport segítségével megkezdte a mutatókészlet felülvizsgálatát annak érdekében, hogy a mutatók szorosabban kapcsolódjanak a megújított SDS céljaihoz és célkitűzéseikhez.

Mit szükséges mérni? Az egyik alapvető kérdés a mutatókészlet körének és céljának meghatározása volt. Mit szükséges mérni: a fenntartható fejlődést magában vagy a fenntartható fejlődési politikákat? Az első esetben kérdés, hogy hogyan lehet a fenntartható fejlődést meghatározni? A Brundtland-formula, azaz „a jelen generáció igényeinek kielégítése a jövő generáció igényeinek veszélyeztetése nélkül”, gyakorlati megfogalmazása köztudottan nehezen alakítható át „működőképes” meghatározássá, és a Brundtland-jelentés a fenntartható fejlődés olyan valóban komplex fo-

² SEC(2005) 161 végleges változat: „A fenntartható fejlődés mutatói az EU Fenntartható Fejlődési Stratégiája megvalósításának nyomon követésére”.

galmát dolgozta ki, mely messze túlmutat egy hangzatos frázison. A Brundtland-jelentés másik, analitikusabb kísérlete a fenntartható fejlődés meghatározására: „egy változási folyamat, melyben az erőforrások kiaknázása, a beruházások iránya, a műszaki fejlődés orientációja és az intézményi változás mind összhangban van, és növeli az emberi igények és törekvések kielégítéséhez szükséges jelen és jövőbeli potenciált”. Ez nyilvánvalóan számos elvet foglal magába, beleértve a generációk közötti és azokon belüli egyenlőséget és igazságosságot, a fenntartott gazdasági növekedést, mely nem károsítja a környezetet vagy teszi tönkre a helyi természeti erőforráskészletet, a szegénység megszüntetését és a döntéshozatalban való részvételt.

Végül az Eurostat válasza – mint sok más országé és szervezeté, mely felállított ilyen mutatórendszert – pragmatikus volt és a politikai prioritásokhoz igazította a mutatókészletet annak érdekében, hogy könnyen meg lehessen határozni a politikaalkotókat a statisztikai munkában. Az ENSZ elgondolásával összhangban szem előtt tartották, hogy a mutatók az EU-SDS alapelvein és célkitűzésein alapuljanak. Továbbá, szintén az ENSZ-modellnek megfelelően, a mutatókat olyan tematikai struktúrába volt célszerű rendezni, mely könnyen érthető a politikaalkotók számára.

A stratégia értelmezhetősége. Szükséges volt ezek után kidolgozni magának az EU-stratégiának a világos értelmezését. A 2001-es SDS prioritási területei, valamint a pótlólagos kormányzati és globális partnerségi fejezetek képezik fő fejezetekként a mutatórendszer gerincét. Két további téma (1. termelési és fogyasztási minták, valamint 2. a gazdasági fejlődés), mely általában átfedi a többi nyolc fejezetet is, szintén külön fejezetként szerepel.

Minden, az említett kötelezettségekből eredő fejezet külön „témává” vált egy olyan hierarchikus szerkezetben, ahol a mutatókat három szintre osztották (lásd az ábrát) azért, hogy illeszkedjenek a különböző típusú felhasználók igényeihez. A 10 téma a következő.³

1. Gazdasági fejlődés.
2. Szegénység és társadalmi kirekesztődés.
3. Öregedő társadalom.
4. Közegészségügy.
5. Éghajlatváltozás és energia.
6. Termelési és fogyasztási minták.
7. Természeti erőforrások kezelése.
8. Közlekedés.
9. Jó kormányzás.
10. Globális kapcsolatok.

³ Az SDI-k felülvizsgálatával kapcsolatos ülésre (2007. május 10–11.) a cikk készítésének időpontjában nem került még sor. Valószínűleg a témák elnevezése bizonyos mértékben változik.

A témákat továbbosztották altémákra és „elintézendő területekre”. Javasolták, hogy a mutatók minden témán belül feleljenek meg a stratégia és a kapcsolódó EU-stratégiák fő politikai kötelezettségeinek és célkitűzéseinek. A mutatókat három fontossági szinten csoportosították egy piramisszerkezetet kialakítva, egy (néhol kettő) fő mutatóval a csúcson, melyet a 2. szintű mutatók (középen) és a 3. szintű mutatók (alul) támasztanak alá. (Lásd az ábrát.)



Az volt a szándék, hogy a piramis különböző rétegei egymáson alapuljanak, és egy-egy történetet mondjanak el. Más szóval, ha valaki tudni szeretné, hogyan egy felsőbb szinten levő mutató miért olyan, amilyen, megkaphatja rá a választ a következő alsóbb szinten. Másik irányból nézve, a piramis egy alsóbb szintjén levő információnak egyenesen kell vezetnie a következő felsőbb szinten levő információhoz. A gyakorlatban ez nem bizonyult lehetségesnek minden téma esetében, különösen a fokozottabban környezeti és intézményi jellegű témáknál.

A mutatók kiválasztásánál szem előtt tartották, hogy az elsődleges cél a nem fenntartható vagy esetleg nem fenntartható tendenciák azonosítása. A mutatókat ezért azon az alapon választották ki, hogy mennyire képesek az EU-trendeket bemutatni, és nem az olyan mutatók mellett döntöttek, melyek inkább az egyes országok relatív teljesítményének összehasonlítására alkalmasak.

A stratégia monitoringja. Bár az Eurostat a fenntartható fejlődés mutatókészleteinek kidolgozásával és közzétételével az 1990-es évek közepétől kezdve foglalkozott, a mutatók a 2006-os stratégiában kaptak új szerepet. Nevezetesen, az Eurostatnak két évente monitoringjelentést kell készítenie a mutatók alapján.

Ez a monitoringjelentés inputot jelent az inkább politikai jellegű fejlődési jelentés számára, melyet az Európai Bizottság fogad el, és amely a mutatókra is hivatkozik. Ez a két jelentés közötti kapcsolat jól láthatóvá teszi a mutatókat, mely csak úgy lehetséges, hogy a mutatók kiválasztási szakaszában konstruktív kapcsolattartás van a

statisztikusok és politikaalkotók között. Ez megköveteli a statisztikai intézmények semlegességéről szóló elv szigorú betartását.

A monitoringjelentés célja, hogy kvantitatív elemzést adjon a különböző fenntartható fejlődési témák trendjeiről. A jelentés szerkezete követi a mutatók keretének 10 témáját. A hangsúly sokkal inkább a trendek grafikonokkal és ábrákkal történő megjelenítésén van, mint a nyers adatokon, melyek elérhetők az Eurostat SDI-weboldalán

(http://epp.eurostat.ec.europa.eu/portal/page?_pageid=2873,63643317,2873_63643793&_dad=portal&_schema=PORTAL). Minden témánál összefoglalják a fő politikai célkitűzéseket a mutatók kiválasztásának magyarázatával. Minden fejezetre egészében és az egyes mutatókra külön elemzik a trendeket. Ez tartalmazza a változások vizuális értékelését: kedvezőek vagy kedvezőtlenek-e a változások a fenntartható fejlődés célkitűzéseinek szempontjából, nem történt jelentős előrehaladás, vagy nem állt rendelkezésre elegendő adat az értékeléshez.

A kapcsolati rész célja, hogy illusztrálja a fenntartható fejlődésre vonatkozó különböző kérdések közötti kapcsolatot. A fenntartható fejlődés kihívása arról szól, hogy ezeknek a kapcsolatoknak jobban tudatosulniuk kell, és azonosítani kell az egyedi tevékenységekben rejlő lehetséges együttműködési lehetőségeket és kompromisszumokat, valamint a kormányok és nemzetközi szervezetek politikáit. Ezek ugyanakkor összetett kérdések, és sokszor korlátozottak a szükséges ismeretek. A kapcsolati rész így nem terjed ki minden területre ebben a szakaszban, inkább néhány hasznos szemléltető példát és hipotézist ad, melyet tovább lehet tesztelni.

Végezetül meg kell említeni a monitoringjelentésben elfogadott megközelítés néhány korlátját. Egyrészt az elfogadott mutatók nem tökéletesek, és nem fedik le mindig megfelelően az adott kérdést. Másrészt az elfogadott mutatók számát tudatosan kommunikációs célokra korlátozták. A jelentésben használt mutatókészlet ezért nem tekinthető kielégítő információnak ahhoz, hogy teljes és átfogó elemzést adjon a fenntartható fejlődéshez kapcsolódó összes tendenciáról. Mindazonáltal a jelentés olyan eszköznek tekinthető a fenntartható fejlődés irányába történő haladás mérésében európai szinten, mely közösen megállapított, főleg a hivatalos statisztikákból származó mutatókon alapul, és így biztosítja azon sok perspektíva egyikét, melyet a még fenntarthatóbb fejlődés keresése megkíván.

Néhány speciális tapasztalat

Semleges értékelés. A statisztikai intézmények különösképpen jó helyzetben vannak ahhoz, hogy alapvető szerepet játsszanak az olyan politikák értékelésében, mint a fenntartható fejlődés, mind a statisztikák pártatlansága iránti elkötelezettségük, mind a számos területen szerzett tapasztalatuk miatt. Az ilyen politikák monitoringja

szempontjából releváns, megfelelő mutatókészletek kidolgozása lehetőséget ad arra, hogy a statisztikákat és a mutatókat fokozottabban használják a nyilvános vitákban. Ennek szükséges előfeltétele nyilvánvalóan az, hogy megtartsák pontosan meghatározott és semleges hozzáállásukat a közérdekű politikák monitoringjában, és különösképpen tartózkodjanak bármilyen politikai nyilatkozattételtől.

A jóváhagyás szükségessége. A legfontosabb érintettek jóváhagyása lényeges tényező a sikeres kimenetel szempontjából. Elsősorban az szükséges, hogy a politikaalkotók elkötelezettek legyenek a mutatók használatára és az általuk esetleg közvetített üzenetek irányába. Ez a közelítési folyamat a mutatók kiválasztási szakaszában kezdődik, amikor is a statisztikusoknak indokolniuk kell preferenciáikat annak érdekében, hogy később a politikaalkotók ne hiúsítsák meg addigi munkájukat.

Az EU-SDI-k esetében 2005 februárjában a Bizottság formálisan jóváhagyta a keretet és a mutatókészletet egy belső közlemény formájában. Ez a Bizottság különböző főigazgatóságainak megkérdezésével és beleegyezésével járt együtt, ami szükségszerűen új tárgyalási körökhöz vezetett arról, hogy mely mutatókat és milyen szinten kell bevonni. A végfelhasználók bevonása ebben a szakaszban alátámasztotta a mutatók, a monitoring és a célok szerepét annak végiggondolásában, hogy hogyan lehet a stratégiát továbbfejleszteni, és garantálta, hogy ezek az elemek bekerülnek a stratégia jövőbeli változataiba.

Végül a munkacsoport zárójelentését jóváhagyásra elküldték a nemzeti statisztikai hivatalok elnökeinek annak érdekében, hogy biztosítsák a nemzeti szakértők további együttműködését és segítségét. Ez a megbízás megújítását eredményezte arra vonatkozóan, hogy egy munkacsoport keretében tovább folytassák a fejlesztést, és biztosította a projekt további támogatását.

Az ütemezés fontossága. Egy projekt esetében alapvető fontosságú a mérőföldkövek meghatározása annak érdekében, hogy az eredményeket a legalkalmasabb időben tudják bemutatni. Különösképpen értékes lehet, ha egy bizonyos politika irányába történő haladás mérésére alkalmazott mutatókat politikai eseményekhez kötnék. Ekkor sokkal nagyobb az esélye, hogy a munkát észreveszik és hasznosítják.

Ezt szem előtt tartva létrehoztak egy külön, a mutatóknak szánt weboldalt még az Európai Tanács 2005. tavaszi ülése előtt, ahol az SDI-stratégiát kellett megvitatni. Később, az Európai Tanács a fenntartható fejlődés stratégiát felülvizsgáló, 2005. decemberi ülését megelőző héten megjelentették az Eurostat jelentését, mely nyomon követte az EU szervezeteinek a fenntartható fejlődés irányába történő előrehaladását.

További figyelmet igénylő területek. A leírt munka egy olyan mutatókészlet elfogadásához vezetett, mely megjelent mind nyomtatott formában, mind az interneten, melyet az Európai Bizottság és az Európai Statisztikai Rendszer jóváhagyott, és mely nagy szerepet kapott az EU fenntartható fejlődési céljának irányába történő

előrehaladás értékelésében. Ugyanakkor fel kell ismerni, hogy további előrelépés szükséges számos területen, melyek közül néhányat bemutatunk.

1. Az SDI-k fogalmi alapja. Lehet érvelni azzal, hogy az EU SDI-mutatókészletének nincs meg a fogalmi alapja vagy kerete, és hogy a kapcsolódása az EU-stratégiához, melyet rendszeresen felülvizsgálunk, azt jelenti, hogy nem fogja kiállni az idők próbáját. Az igaz, hogy az alkalmazott strukturális keret nagymértékben az EU-stratégiából származik, és így a stratégia alakulásával változásnak van kitéve. A mutatókészlet szerkezete ugyanakkor világos és könnyen kommunikálható. Olyan eszközt biztosít, mely segíti a változás végrehajtását a stratégia célkitűzései felé történő előrehaladás mérésével, és amelyet már ténylegesen alkalmaztak az EU döntéshozatalában. Várható, hogy a stratégia alakulásával új mutatókra lesz szükség, melyek az egyre növekvő súlyú kérdésekre vonatkoznak. Egy közös UNECE/OECD/Eurostat-munkacsoport jelenleg is foglalkozik a legfőbb megközelítés részleteivel, melyhez még multidiszciplináris kutatásra van szükség a lehetséges megvalósítás előtt. Szükséges annak bizonyítása is az érintettek felé, hogy egy ilyen keret hasznos a jelenlegi fenntartható fejlődés politikák monitoringja szempontjából.

2. Nem kielégítő adatok. Több olyan terület is van, amelyek vonatkozásában jelenleg nincsenek információk vagy csak részleges információk állnak rendelkezésre. Ennek a technikai akadálynak legyőzése és a politikaalkotáshoz hosszabb távon szükséges adatok előállítására és összeállítására érdekében a mutatókat két kategóriába osztották, a „leginkább rendelkezésre álló” és a „kifejlesztendő” mutatókra. A „leginkább rendelkezésre álló” mutatók azok a mutatók, melyek a létező adatok alapján összeállíthatók. Néhány ilyen mutató lehet, hogy nem a legideálisabb a fenntartható fejlődés kérdései szempontjából, de hasznos szerepet tölthetnek be a még „kifejlesztendő” ideális mutatók proxyjaiként. Néhány mutató adatminőségi problémák miatt szintén a „kifejlesztendő” kategóriába sorolandó. Az Eurostat az európai statisztikai rendszerrel és a kutatóközösséggel együttműködve megvizsgálta a „kifejlesztendő” mutatók alkalmazhatóságát, és jelentést készít az alkalmazható mutatók rendelkezésre állásának alakulásáról. Ez a fejlesztő munka jelentős feladat, mely alapvetően hozzájárul a mutatókészlet homogenitásának további javításához.

3. Generációk közötti egyenlőség. Az SDS tartalmaz olyan tevékenységeket és kötelezettségeket, melyeket most kell teljesíteni a nem fenntartható tendenciák korigálása érdekében, és melyek célja a jövő generációk helyzetének és lehetőségeinek javítása is. A generációk közötti egyenlőség így alapvető fontosságú része a stratégiának. Előrejelzés és forráskönyv-építés szükséges a statisztikai információk relevanciájának javításához, de ezeket az eszközöket még ki kell alakítani.

4. Tendenciák értékelése. Az egyes mutatók tendenciáinak eddig alkalmazott értékelési módszere harmonizált ugyan, de túlságosan leegyszerűsített. Mindegyik mutató úgy értelmezhető, hogy emelkedő, csökkenő vagy stabil-e a fenntartható fejlődéshez kapcsolható tendencia. De a mutató milyen mértékű mozgása szükséges a he-

lyes irányba, hogy fenntarthatónak lehessen tekinteni? Az Eurostat monitoringjelentésében elfogadott módszer pozitívnak értékeli a változást, ha az a helyes irányban éves átlagban legalább 1 százalék több éven keresztül. Ez egy egyszerű, következetes és átlátható gyakorlati szabály, de természetesen nem kielégítő olyan szempontból, hogy nem tud differenciálni az olyan változékonyabb változókat, mint például a GDP növekedési rátáját, és a kevésbé változékony változókat, mint például a várható élettartam között. Ahol a mutatók mennyiségileg meghatározott célokhoz köthetők, az értékelés a szerint történik, hogy a fejlődés megegyezik-e a lineáris célkitűzés-menettel. Ha azonban maga a cél nem alkalmas erre, ez a módszer nem alkalmazható. Az értékelési munka világossá tette, hogy szükséges néhány specifikus (lehetőleg fenntarthatósági) küszöböt meghatározni minden mutatóra. Szükséges a mutatókészlet általános trendjeit összességében is értékelni, azért, hogy látni lehessen, hogy a fenntartható vagy a nem fenntartható úton haladunk-e.

5. *A kormányzás mutatói.* A kormányzás kérdéskörének lefedése problematikus téma. Ideális esetben a fenntartható fejlődéssel összefüggésben a mutatóknak az intézményekre összpontosítva le kellene fedniük az olyan területeket, mint a részvétel, elszámoltathatóság, hatékonyság és koherencia, és érinteniük kellene az EU szerepét is a globális kormányzásban. Tekintettel a megfelelő, kielégítő minőségű és körű adatok megszerzésének nehézségeire csak két altémát (politikai koherencia és társadalmi részvétel) és néhány nem tökéletes mutatót javasoltak a témában.

Következtetések

A fenntartható fejlődés fogalmának bevezetése a jelenlegi politikaalkotásba fontos fordulópontot jelentett társadalmaink számára az utóbbi két évtizedben. A jelenlegi politikák nem koncentrálhatnak már elsősorban csak rövid távú hatásaikra, hanem előrébb kell tekinteniük, és egymás között is nagyobb összhangot kell teremteniük. Tekintettel a fenntartható fejlődés fogalmának – mely azt méri, hogy mi számít a jelen és a jövő generáció jólétének – összetettségére, a hivatalos statisztikák megfelelően összeállítottak ahhoz, hogy mind a pártatlanság iránti elkötelezettségük, mind a rendelkezésre álló tapasztalatok sokszínűsége miatt olyan szilárd statisztikai eszközöket – elsősorban statisztikai mutatószámokat – biztosítsanak, melyek szükségesek a jelenlegi politikák megvalósításának megfelelő értékeléséhez. Az EU jelenlegi fenntartható fejlődési mutatókészlete ugyan még mindig jócskán hiányos, de egy olyan előrejelző megközelítés, melyet az Eurostat követ, bővíti a hivatalos statisztikák profilját, és hozzájárulhat ahhoz, hogy a jövőbeli politikákat a jelenlegi helyzet pontosabb értékelésén alapulva alakítsák ki.

Hírek, események

Jutalom. Közzolgálati jogviszonyban töltött idejük alapján 2008. március hónapban jubileumi jutalomban részesültek: 25 éves szolgálatért: *Sugár Emese*, Árstatisztikai főosztály; *Nagyné Pakula Urszula*, Tájékoztatási főosztály; *Kavecsánszkiné Bosnyák Mária*, Miskolci Igazgatóság; *Kolozsvári Zsuzsanna*, Statisztikai Elnökhelyettesi Titkárság; *Dobránszkiné Óvári Klára*, Vállalkozásstatisztikai főosztály. 30 éves szolgálatért: *Szénásiné Matúz Ágnes*, Gazdálkodási főosztály; *Czukender Gyula*, Tájékoztatási főosztály; *Gyulai Barnabás*, Gazdálkodási főosztály; *Madarász Gyöngyi*, Informatikai főosztály; *Szűcs Lászlóné*, Miskolci Igazgatóság; *Tüske István*, Árstatisztikai főosztály; *Ercsey Zsófia*, Statisztikai kutatási és módszertani főosztály. 35 éves szolgálatért: *Dr. Horváth József*, Pécsi Igazgatóság; *Meskó Anna*, Nemzeti számlák főosztály; *Farkas Veronika*, Szegedi Igazgatóság; *Laklott Györgyi*, Igazgatási és nemzetközi főosztály.

Az Eurostat Statisztikai Program Bizottságának (SPB) 65. ülését és a „Statisztikai kihívások” munkacsoportjának 5. ülését 2008. február 14. és 15. között tartották Luxembourgban. A Statisztikai Program Bizottság 65. ülésén két döntést igénylő napirendi pontról tanácskoztak a szakemberek a külföldi tulajdonú vállalkozások statisztikájának tárgyában. Vitára bocsátott kérdések voltak még a nemzeti számlák, a harmonizált fogyasztói árindex, a szezonális kiigazítás, valamint az ENSZ Statisztikai Bizottság ülésének napirendje. Az SPB két napirendi pontban döntést hozott, négy napirendi pontot megvitattak a

résztevők, további hét témában pedig tájékoztatást kaptak az SPB tagjai.

A „Statisztikai kihívások” munkacsoport az előzetesen kiküldött anyagok, valamint a tagok által írt és előzetesen köröztetett dokumentumok alapján dolgozott. A főbb irányokat az előző megbeszélések alkalmával a tagok meghatározták, az ülésen a résztvevők megvitatották a társadalomstatisztikával kapcsolatos stratégiai kérdéseket, a hivatalos statisztika szerepének erősítésére vonatkozó lehetőségeket, a módszertani alcsoport további működésének irányát, az Európai Központi Bankkal folytatott megbeszélés eredményeit, valamint a 2008-ban, Vilniusban tartandó DGINS-konferencia előzetes kérdéseit. A megbeszélésen *dr. Pukli Péter*, a Központi Statisztikai Hivatal elnöke vett részt.

Az UNECE–OECD–Eurostat Fenntartható fejlődés statisztikája munkacsoportja Lisszabonban 2008. március 5. és 6. között megtartotta ötödik közös értekezletét. A munkacsoport az elmúlt két évben megrendezett három értekezleten és annak eredményeképpen elvégzett munka során elkészítette azt a jelentést, amelynek célkitűzése egy olyan referenciaanyag kibocsátása, az egyes nemzetek statisztikai hivatalai számára, amely egységes fogalmi keret biztosításával alkalmas lehet a fenntartható fejlődés mérésére. A lisszaboni értekezleten jelentésbe foglalták az előző, bukaresti értekezleten felmerült hozzászólásokat, véleményeket a végső jelentést elfogadták. A munkaterv szerint a jelentés véglegesítésének határideje 2008 tavasza, mivel 2008 júniusában azt az Európai Statisztikai

tikusok Plenáris Ülése elé kell elfogadtatás céljából beterjeszteni.

Az értekezleten egyetértés született a hivatalos statisztikák használatának szükségességéről, valamint az indikátorok közzétételének, népszerűsítésének fontosságáról. Magyarország számára fontos, hogy e téren helyet kapjanak országunk és térségünk szempontjai. A munkacsoport eredményei alapját képezhetik egy jövőbeni hazai indikátorfejlesztési munkának. A tanácskozáson *Mészáros Andrea*, a KSH tanácsosa vett részt.

Az Országos Statisztikai Tanács 2008. március 20-án tartotta ülését a KSH-ban. Az ülés napirendje a következő volt. 1. Tájékoztató az Európai Statisztikai Törvény felülvizsgálatáról, az európai statisztikai rendszer várható struktúrájáról és irányítási viszonyairól. Az előterjesztők: *dr. Nagy Eszter*, a KSH tanácsosa és *dr. Láng Dóra*, a KSH osztályvezetője voltak. 2. Tájékoztató a 2011. évi népszámlálással összefüggő kérdésekről. A napirendi pont előterjesztője *Waffenschmidt Jánosné*, a KSH főosztályvezetője volt. 3. Az osztályozások változása 2008-tól. A napirendi pont előterjesztője *Rónainé dr. Györgyi Márta*, a KSH szakmai tanácsadója volt. 4. Egyebek: a) Tájékoztató az OST Regiszter Munkabizottsága és a 2212/2007. (XI. 14.) Korm. határozattal létrehozott munkacsoport együttes munkájáról. Az előterjesztő *Gárdos Éva*, a KSH statisztikai tanácsadója volt. b) Az európai statisztikai rendszerhez tartozó magyar szervezetek szakmai önértékelésének eredménye. Az előterjesztő *Mészáros Árpád*, a KSH főosztályvezetője volt.

A KSH Vállalkozás-statisztikai Szakmai Kollégiuma 2008. március 18-án tartotta

ülését. Az értekezlet napirendjén az új ipari termékosztályozás (ITO) és az új mezőgazdasági termékosztályozás (METO) bevezetésének kérdései szerepeltek. *Kotulics Tamás*, a KSH főosztályvezetője, a szakmai kollégium vezetője érzékeltette, hogy folytatódik a nemzetközi nomenklatúrák átalakulása, melyet szükségszerűen a magyar statisztikai nomenklatúrák átalakulása, továbbfejlesztése is követ. Ebbe a sorba tartoznak a kollégium elé kerülő új ipari termékosztályozás (ITO) és az új mezőgazdasági termékosztályozás (METO) tervezetei. Elmondta, hogy a kollégiumi előterjesztések jelentős előkészítési munkákra alapozódtak, melyek a korábbi hónapokban szakértői szinten folytak. A szakértői egyeztetések eredményeképpen a jelen kollégiumi ülés előkészítésekor lebonyolított hivatali egyeztetés során egyetértés és támogatás volt tapasztalható, az észrevételek nagy része technikai jellegű volt. „Az új ipari termékosztályozás (ITO) bevezetése” napirendi pont előterjesztője *Csorba László*, a KSH osztályvezetője, „Az új mezőgazdasági termékosztályozás (METO) bevezetése” napirendi pont előterjesztője *Pintér László*, a KSH főosztályvezető-helyettese volt.

Sajtóbeszélgetések. *Dr. Laczka Éva*, a KSH főosztályvezetője és *Végh Zoltán*, a KSH Szegedi Igazgatóságának igazgatója a gyümölcsös ültetvények összeírásáról 2008. március 4-én sajtóbeszélgetés tartott a KSH-ban. A Központi Statisztikai Hivatal Fényes Elek termében pedig 2008. március 18-án tartottak sajtóbeszélgetést, amelynek témája a 2007. évi népmozgalmi folyamatok demográfiai összefüggései volt. A beszélgetés előadója *Kamarás Ferenc* demográfus, a KSH főtanácsadója volt.

**A Nemzetközi Statisztikai Intézet (International Statistical Institute – ISI)
fontosabb konferenciaajánlatai**

(A teljes ajánlatlista megtalálható a <http://isi.cbs.nl/calendar> honlapon.)

Berlin, Németország. 2008. május 15–16.

A Német Statisztikai Társaság pünkösdi találkozója (The *Pentecost Meeting of the German Statistical Society*.)

Információ: *Nana Dyckerhoff* (Managing Director, German Statistical Society)

Telefon: +492214704130

Fax: +492214705084

E-mail: post@dstatg.de

Honlap: www.dstatg.de

Sousse, Tunézia. 2008. június 2–6.

Nemzetközi konferencia a random mátrixokról. (International Conference on Random Matrices „ICRAM”.)

Információ: *Abdelhamid Hassairi*

Telefon: (216) 98 667 564

E-mail: abdelhamid.hassairi@fss.rnu.tn

Honlap:

www.tunss.net/accueil.php?id=ICRAM

Sozopol, Bulgária. 2008. június 21–28.

XIII. Nemzetközi nyári konferencia a valószínűségről és a statisztikáról és Szeminárium az adatelemzésről. (XIII International Summer Conference on Probability and Statistics and Seminar on Statistical Data Analysis.)

Információ és jelentkezés online: *Nick Yanev*

Telefon: +35929793861

Fax: +35929713649

E-mail: yanev@math.bas.bg,
statlab@math.bas.bg

Honlap: <http://stochastics.fmi.uni-sofia.bg>

Dubrovnik, Horvátország. 2008. június 23–26.

BIOSTAT 2008. Statisztikusok és biometriai kutatók 15. találkozója együtt az ITI 2008-al, a 30. Nemzetközi konferenciával az információs technológiai interfészeiről. (BIOSTAT 2008 – 15th Meeting of Researchers in Biometrics/Statistics in conjunction with ITI 2008 – 30th International Conference on Information Technology Interfaces.)

Információ: Conference Secretariat – ITI 2008 E – University Computing Centre, University of Zagreb J. Marohnica 5, HR-10000 Zagreb, Croatia

Telefon: +385 1 616 5597

Fax: +385 1 616 5591

E-mail: iti@e.hr

Honlap:

<http://iti.e.hr/html/biostat2008.html>

KÖNYVSZEMLE

Dr. Nemes Erzsébet (szerk.):

Fényes Elek breviárium

KSH Könyvtár és Levéltár. Budapest. 2007. 244 old.

A Központi Statisztikai Hivatal és a Magyar Statisztikai Társaság azzal tiszteleg *Fényes Elek* emléke előtt, hogy születésének 200. évfordulója alakalmából reprezentatív áttekintést ad sokrétű, gazdag munkásságáról. Az előttünk levő kötet *dr. Nemes Erzsébet* a KSH Könyvtár és Dokumentációs Szolgálat főigazgatója körültekintő, hozzáértő munkájának eredménye, aki elismerésre méltóan jól válogatott műveiből illetve azok részleteiből.

A kötet átfogó képet ad Fényes Elek munkásságáról, s azon keresztül a reformkori Magyarország társadalmi és gazdasági viszonyairól, politikai csatározásairól. A jól megválasztott szemelvények mellet közreadott fotók képileg is érzékletessé teszik a kor hangulatát, stílusát.

Fényes Elek forradalom előtti és utáni munkássága jól megkülönböztethető. A kötet az 1836 és 1849 közötti időszakot bemutató I. része a legelsőként megjelent, már említett célt megvalósító hat kötetes sorozat a „Magyar országnak, 's a' hozzá kapcsolt tartományoknak mostani állapotja statistikai és geographiai tekintetben” kezdődik.

Az első kötet előszavából kiderül, forrásként *Lipszki János* 1830-ra vonatkozó adatait használta. Nagy gondot jelentett a népesség számának megállapítása, már csak azért is, mert az összeírásokból a nemességet kihagy-

ták. Adatokat közül a népesség vallási és nyelvi-nemzetiségi megoszlásáról is, jöllehet tudja, hogy adatai meglehetősen bizonytalanok.

Az egyes kötetek foglalkoznak a vármegyék földrajzi fekvésével, határaival, nagyságával, természeti tulajdonságaival, hegyeivel, folyóvizekkel, tavakkal-mocsarakkal, ásványos és orvosi vizeivel, termékeivel (növények, állatok, ásványok), lakhelyekkel és lakosokkal, a „vármegye szorgalmával”, ami az ipari és kereskedelmi tevékenységet jelentette. Adatok találhatóak a közintézményekről, iskolákról, a katonaságról, a jobbágytelkekről, nemesi uradalmakról valamint a megyék politikai elosztásáról. Elsőként a szabad királyi városokat említi, majd járások szerint a településeket, részletezve a mezővárosokat, népes pusztákat stb. Tájékoztat a falut körülvevő föld minőségéről, a terményekről, a földesurakról, név szerint említve őket. A megyék felsorolása után a hatodik kötet a Magyar katonai végvidék leírását tartalmazza.

Minden kötetből egy-egy vármegyét kiemelve a breviárium érzékelteti Fényes Elek rendkívül alapos munkáját.

Érdekes és hasznos a hatodik kötet végén található mutató, ami nemcsak a települések kikereshetőségét biztosítja, hanem utalásokat tesz a gazdaságra, kereskedelemre, közigazgatásra, népességre stb.

A következő mű a „Magyarország statistikája” logikus folytatása az első nagy műnek, amelyben a geográfiai leíráson túllépve az ország „alaperejét” (ipar, kereskedelem stb.) valamint az ország polgári szerkezetét és kormányzását is bemutatja.

Az első kötet igazi statisztika. Igyekszik adatforrásait bővíteni, bizonyos kritikával illeti

azokat, jelzi, hogy ez a munka egy kezdet, amit további, egyre jobb összeállítások fognak követni. Az adatok bemutatásában is törekszik az egyszerűsége, (kerüli például a nagyméretű táblázatokat) amint ma mondanánk „felhasználóbarát” prezentációra törekszik. A második kötet a gazdasági és társadalmi élet jogi hátterét, az alaptörvényt mutatja be. A harmadik kötet a közjövödelmekkel, a tudományos és oktató intézetekkel, az igazságüggyel, az egyházakkal és a „politiá”-val foglalkozik, ami egyet jelentett a mindennapok rendjével, az egészségügyben, a szociális ellátásban, a pénzügyekben. Hangsúlyozza a takarékpénztárak szerepét, ami jótékony hatással van a szegények erkölcsére, megtakarítási hajlandóságára. Foglalkozik a takarékpénztárak működésének nemzetközi összehasonlításával is. Ebben a munkában már megjelenik Fényes Elek igazi statisztikus énjé, amennyiben a rendelkezésére álló eszközöket a minél jobb statisztikai adatok előállításának szolgálatába állítja.

A breviárium harmadikként az 1844-ben Lipcsében megjelent „Magyarország hátramaradása ügyében felelet dr. Wildner Ignác urnak Fényes Elektől” c. vitairatát mutatja be. Wildner osztrák jogtudós írásban bírálja a magyar gazdaságot, a korszak mértékadó lapjában, az *Augsburger Allgemeine Zeitungban* az államkincstár ürességének alapvető okát, a nemesség adómentességét, de szóba kerül az ország és a birodalom viszonya is, amennyiben, a kritika szerint, Magyarország nem járul hozzá eléggé sem pénzügyileg, sem katonailag az osztrák birodalom fenntartásához.

A négy szakaszba rendezett könyv egyfajta párbeszéd az ország korabeli állapotairól ért vádak és Fényes Elek több száz évre visszamenően idéz fel történelmi eseményeket, vitatkozik, érvel, majd arra a következtetésre jut: ... „a’ magyar birodalom katonasága mind számra, mind vitézségre nézve a’ többi ausztriai nem magyar tartományok hadával

tökéletesen arányban áll”. A szellemi és anyagi javakat azonban pótolni kell és kiáll amellett, hogy a nemességnek is ki kell vennie ebből a részét.

1847-ben jelent meg két kötetben a „Magyarország leírása”. Mint a bevezetőben írja, eleget tesz barátai kérésének, hogy az addig kilenc kötetben megjelent földrajzi és statisztikai tárgyú műveiből könnyebben áttekinthető összefoglalót készítsen. Az adatok nagy részét felfrissítette, mivel, a „statistica állandóan változik”. Említést tett arról is, hogy tervezi az ország iparát feltáró statisztika összeállítását, valamint egy geográfiai nagyszótár megszerkesztését.

Az ereje, alkotókedve teljében levő Fényes Elek hozzáfogott a tervezett nagy mű megírásához. Ez a „A magyar birodalom statistikai, geographiai és történelmi tekintetében”, melynek alcíme: „Részletes és kimerítő leírása Magyar és Erdélyországnak mely külön fejezetben terjeszt elő minden magyar és erdélyországi vármegyét, kerületet, vidéket, székot; s ezekben járasonként ismét betűrendben iratik le minden királyi és mezőváros, falu és pusztá; s minden egyes helynek nemcsak fekvése, népessége nyelv és vallás szerint, termékenysége, földesura; hanem hol csak lehetett határki-terjedése (mennyi belőle úrbéri, s mennyi allodiális, mennyi szántóföld, rét erdő, szőlő, sat.) sőt nevezetes helyek története is előadatik”. E hosszú, tartami leírásból álló alcím önmagáért beszél. Jelzi, hogy a szerző ezt legnagyobb, legsokoldalúbb művének szánta.

1848-ban megjelent a Komárom megyével foglalkozó első kötet. Érzékelhető volt, hogy milyen nagy előrelépést jelentett az adatok sokrétűségében, pontosságában. Ez az év azonban sorforduló volt Fényes Elek számára. A forradalom, hazafias elkötelezettsége gyökeres változást hozott szakmai munkásságában is. Az elkezdett munka megszakadt, terveit nem tudta valóra váltani.

A breviárium II. része Fényes Elek forradalom és szabadságharc utáni statisztikai tevékenységét ismerteti.

A „Magyar ország geographiai szótára, mellyben minden város, falu és pusztá, betűrendben körülményesen leiratik, négy kötetben” 1851-ben jelent meg. Könyvtároszakmai szemmel tekintve Fényes Elek életművének legidőállóbb munkája, olvasható a II. rész bevezető tanulmányában. Igazi helység-névtárral van dolgunk, mely azonban nem előzmény nélküli. Szorgalmas, folyamatos adatgyűjtő munkája, korábbi tapasztalatai, kapcsolatai segítették.

Fényes Elek megállás nélkül dolgozik. 1854-ben közreadja „A török birodalom leírása, történeti, statistikai, geographiai tekintetben” című munkáját. Sok érdekes adat található a kötetben számos magyar történelmi vonatkozással. Ugyancsak ebben az évben jelent meg: „Az orosz-török háború történeti jegyzetekkel, okiratokkal, szereplő jelesb férfiak életrajzával, földabroszokkal és ütközeti térképekkel felvilágosítva”. Az, hogy ebben az időben Fényes érdeklődése a külpolitika irányába fordult, bizonyos fordulópontot, törést jelentett. Már érezhető, hogy a forradalom leverése utáni, jelentősen megváltozott környezetben számára új helyzet állt elő. Ebbe a sorba illeszkedik az 1857-ben megjelent „Az ausztriai birodalom statistikája és földrajzi leírása” című munkája, amely jórészt osztrák forrásokra támaszkodott. Tudjuk, ebben az időben, már megélhető gondjai vannak, ezért munkájában az anyagi szempontok is fontos szerepet játszanak.

A következőként bemutatott munkájának címe sokat ígér, de csak egy kötet jelent meg belőle 1859-ben. „Magyarország 1859-ben statistikai, birtokviszonyi és topographiai szempontból. Részletes és kimerítő leírása az egyes vármegyéknek, s azokban a 100 holdnál többet bíró birtokosoknak, s szolgabírói járás-

sok szerint az egyes városoknak, mezővárosoknak, faluknak és pusztáknak, az utolsó posta feljegyzésével”.

„A magyar elem s ellenesei” c. 1860-ban megjelent írása valójában röpirat, reagálás egy sajtókampányra, ami akörül bontakozott ki, hogy tekinthető-e Magyarország minden állandó lakosa magyarnak. Fényes Elek szenvedélyesen harcolt az ország nemzetiségi megosztása ellen, nem tagadva, hogy vannak nemzetiségek, de azok szokásaikban (például temetkezés) a magyarokat követik.

Az 1860-as évek elején elérkezettnek látta az időt, hogy a haza jobb megismerése céljából megújítsa Magyarország statisztikáját, hivatkozva a harminc évvel korábbi hat kötetes munkájára. „Magyarország ismertetése statistikai, földirati s történelmi szempontból” című művét négy kötetre tervezte. Az első kötetet két részben 1865 és 1866-ban jelentette meg, amelyek tizenegy dunántúli megye leírását tartalmazta egységes szerkezetben: általános statisztikai rajz, kiterjedés, természeti leírás, folyóvizek, tavak, mocsarak, növényi, állati és ásványi termékek, lakhelyek népessége vallás és nemzetiség szerint, oktatásügy, ipar és kereskedelem, közlekedés és posta, egyletek, egészségügy, árvaellátás, uradalmak, tisztikar, adóügy, közigazgatás és országgyűlési képviselő. Lenyűgözően sokrétű ez a felsorolás különös tekintettel arra, hogy nem volt mögötte hivatal, apparátus. Sajnos azonban az első kötetnél nem jutott tovább.

A breviárium II. részének végén Fényes Elek 1867 áprilisában, a Magyar Tudomány Akadémián elhangzott előadásának könyvváltozata található. Ez is vitairat, melyben azokkal az osztrák centrista vádakkal vitatkozik, miszerint a magyarok az országban kisebbségben vannak, de ez a kisebbség nyelvét, intézményeit rákényszeríti a többi nemzetiségre. Érdeemes emlékeztetni arra, hogy a dolgozat a nemzetiségeket három csoportba sorolja: egy-

millió lélekszám feletti: a magyar, a román, a tót és a német; ennél kevesebben, de százezer-nél többen vannak a szerbek, a rutének és a horvátok, míg a sokácokat, vindusokat (vendek), bolgárokat, olaszokat, görögöket és a klementínusokat az „apró nemzetiségekhez” sorolja. Nem foglalkozott a zsidókkal, az örményekkel, a cigányokkal és a franciákkal, mivel ezek valamely már felsorolt nemzetiség nyelvét használták. Adatai főként az 1857. évi népszámlálásból származtak, de azokat kiegészítette a vallásfelekezeti névtárakból.

Fényes Elek aktív szerepet játszott a forradalom előkészítésében, a kibontakozó reformmozgalom vezéralakja, számos egyesület és a nemzeti radikális kör tagja, újságot ír szerkeszt (Magyar Gazda). Cikkeiben foglalkozott a magyar fölművelés helyzetével, az adó és pénzügyi rendszer kérdéseivel. A forradalom bukása utáni években is tevékenyen publikál, harcol az ország függetlenségéért, a nemzetiségi arányok torzítása ellen

A kötet III. része, melyet Szemelvények Fényes Elek időszaki kiadványokban és szerkesztésében megjelent műveiből címmel állított össze a szerkesztő, jó keresztmetszetet adja politikai, közéleti tevékenységének. Szemelvények találhatóak a mezőgazdasággal foglalkozó írásaiból. Összeállítja a „Magyar gazdasági névtár”-t. Sokoldalúságát jelzi, hogy részletesen elemzi a pénzügyi problémákat „A magyarországi adó- és finanszrendszer 1849-dik évig” című munkájában.

A forradalom leverése utáni helyzetben is sokat foglalkozik statisztikával, de mint láttuk, a visszhang már nem a régi. Dolgozik a biztosítás területén (élettartam és halálozási valószínűségi táblákat készít), belebonyolódik az al-dunai konföderáció vitájába stb.

Kiemeli a breviárium a Statisztikai és Nemzetgazdasági Közlemények 1867. évi 3. kötetéből „Párhuzam egyfelől a magyar koronai birodalom, másfelől az ausztriai német,

lengyel és cseh koronai országok közt, területi, népességi, had- és pénzügyi tekintetben” című írását, amely igazi összehasonlító statisztika. Utolsónak bemutatott cikksorozata a *Magyar Újságban* jelent meg az „Észrevételek Magyarország 1870-ki népességének táblázataira, melyeket az országos stitistikai hivatal közzétett” címmel. Kritikája az 1857. évi és az 1870. évi népszámlálás adatai közötti ellentmondásokat tárja fel.

A breviárium zárófejezete a IV. rész részletes bibliográfiát tartalmaz Fényes Elek műveiből.

*

Fényes Elek nem elfeledett, elhallgatott személyisége történelmünknek. Emlékét szülőfalujában, Csokajban is őrzik. 2001-ben, halálának 125. évfordulóján szülőfalujában a Fényes Elek Társaság Fényes Elek Díjjal tüntetik ki azokat, akik Erdélyben ápolják a honismeretet.

Összefoglalva: a breviárium jó áttekintést ad Fényes Elek írásban fennmaradt életművéről. Érzékelteti korának szellemiségét, stílusát, amit jól illusztrálnak a címlapfotók, valamint a szövegközi fotóillusztrációk.

A kutatóknak és érdeklődőknek segítséget nyújt a pontos bibliográfia, a KSH-könyvtárbeli helymeghatározás. Dicséretes a kötet tipográfiai igényessége, szép kivitelezése. Különösen ajánlható olvasmány azoknak, akik tárgyyszerű képet akarnak szerezni Magyarország nemzetiségi viszonyainak történetéről, az azzal kapcsolatos vitákról, sokszor magyarellenes szándékokról.

Dicséretes a könyv tipográfiája, elegáns megjelenítése, amit gazdagítanak az egyes művek címlapjának fényképei is.

Marton Ádám

kandidátus, a Központi Statisztikai Hivatal ny. osztályvezetője
E-mail: adam.marton@ksh.hu

Folyóiratszemele

**Clements, K.W. – Izan, I. H. Y. –
Selvanathan, E. A.:**

Sztocasztikus indexszámok: áttekintés

(Stochastic index numbers: a review.) –
International Statistical Review. 2006. évi 2. sz.
235–270. old.

Az indexelméletnek két fő iránya van. Az első az ún. próbákön alapuló megközelítés (test approach), a másik a hasznossági függvényeket használó közgazdasági indexelmélet. Kevésbé ismert, de az utóbbi időben egyre inkább az érdeklődés előterébe került az ún. sztocasztikus felfogás (stochastic approach – SA). Az árváltozások esetében ez azt jelenti, hogy az inflációs ráta a folyamat valamiféle alapvető paramétere, trendje (underlying vagy true), amit az egyedi árak mozgásából lehet becsülni. A megfigyelt egyedi adatok (árak) azonban „zajosak” (noisy). A kérdés az, hogyan becsüljük, hogyan nyerjük ki (signal-extraction) a megfigyelt adatokból ezt az ismeretlen paramétert, az inflációs rátát úgy, hogy a mérési hiba hatását minimálisra csökkentjük. Bizonyos esetekben ez a módszer is a jólismert formulákhoz vezet, mint például Divisia-, Laspeyres-, Fisher- stb. indexek, de a bizonytalanság az SA esetében kulcsszerepet játszik, így ez a felfogás különbözik a hagyományos determinisztikus megközelítéstől. Az SA nemcsak a pontbecslést adja meg, hanem annak szórását is, azt, hogy az egyedi ármozgások mennyiben térnek el a közös trendtől, azaz hogy az arányok mennyiben változtak. Ebből következik, hogy az infláció mértéke annál „bizonytalanabb”, minél nagyobbak a relatív árváltozások.

Az SA gondolata már *Jevons* és *Edgeworth* munkáiban [1863, 1887] megtalálható. Később – lehet, hogy *Keynes* kritikája nyomán – feledésbe merült. Az SA új életre kelt *Selvanathan* és *Prasada Rao*, *Diewert* és mások munkássága nyomán. (*Diewert, W. E.* [1995]: *On the stochastic approach to index numbers*. Discussion Paper. 95–31. old. Department of Economics. University of British Columbia.)

A sztocasztikus árindex modelljének lényege röviden a következőképpen vázolható.

Egy termék árindexének logaritmus (Dp_{it}) a következőképpen értelmezhető:

$$Dp_{it} = \alpha_t + \varepsilon_{it},$$

ahol $Dp_{it} = \log p_{it} - \log p_{it-1}$, ($i=1,2,\dots,n$)

α_t az árváltozások minden termék esetében érvényesülő általános trendjét, ε_{it} a véletlen tényezőt (disturbance) jelzi. $E(\varepsilon_{it})=0$. Dp_{it} várható értéke $E(Dp_{it})$ pedig megadja az árváltozások általános trendjét, az infláció mértékét. Az ε_{it} úgy is felfogható, mint az i . termék relatív árának változása (az árváltozás eltérése az átlagtól).

Az egyes termékek súlya a fogyasztásban (a termelésben, a kereskedelemben stb.) különböző. Ezért az árindexek súlyozott átlagok. A levezetés részleteit mellőzve, de az egyes termékek fogyasztásban játszott súlyát is figyelembe véve (értelmezve azt, mint a t , vagy a $t-1$, időszak súlyaiként, vagy azok átlagaként) vagy az infláció mértékének logaritmus a következő formulával becsülhető:

$$\hat{\alpha}_t = \sum_{i=1}^n w_{it} Dp_{it} .$$

Megjegyzés. A *Folyóiratszemelet* a KSH Könyvtár (Orbán-Szirbucz Zsófia) állítja össze.

Ez az index végül is egy Divisia-, vagy Törnqvist–Theil-árindex, amelyeknek – mint az ismeretes – sok jó tulajdonságuk van. (Hulten, C.R. [1973]: Divisia index numbers. *Econometrica*. 41. évf. 6. sz. 1017–1025. old.)

A következő, ún. Divisia-variancia, a termékek keresztmetszeti árány-változásainak szórását jelzi:

$$\Pi_t = \sum_{i=1}^n w_{it} (Dp_{it} - DP_t)^2,$$

ahol $DP_t = \sum_{i=1}^n w_{it} Dp_{it}$ a Divisia-index.

A mintavételi hiba négyzetének becslése n elemű minta esetén:

$$\text{var } \hat{\alpha}_t = \frac{1}{n-1} \Pi_t.$$

Ez azt jelenti, hogy az infláció becslésének mintavételi hibája annál nagyobb, minél nagyobb a relatív árváltozások szóródása.

A felvázolt eredmények több irányban általánosíthatók. Amennyiben az infláció folyamatában bizonyos termékcsoportok között strukturális árányváltozás következik be, akkor azok elkülöníthetők termékspecifikus paraméterekkel:

$$Dp_{it} = \alpha_i + \beta_i + \varepsilon_{it}$$

A β_i paraméter lehetővé teszi a „teljes” infláció termékcsoportonkénti összetevőinek vizsgálatát.

Az egyes termékcsoportokon belül az inflációs folyamatok, azokat önálló egységként kezelve, a már elmondottak szerint elemezhetők. (Árindexekről általában időbeni változások esetén beszélünk. Az elmondottak értelemszerűen, területi, országok közötti árszínvonal-eltérések mérésénél is használhatók, illetve érvényesek a volumenindexek esetében is.)

A szakirodalom szerint az egyedi árindexek súlyozatlan számtani átlaga (Carli [1764]) vagy azok egyszerű geometriai átlaga (Jevons [1865]) voltak az első sztochasztikus árindexek, mivel azok ajánlásában kifejtett szempontok valójában az itt elmondottakhoz hasonló elmélet megfogalmazását jelentették. Edgewort [1888] és Bowley [1901] munkássága továbbfejlesztette az elméleti kereteket, amelyben fontos szerepet játszott a becslés hibájának (az egyedi árindexek szórásának) kiszámítása is, amiből az átlagos árindex (infláció) becslésének megbízhatóságára, stabilitására lehetett következtetni.

Jevons az arany értékcsökkenését akarta mérni. Abból indult ki, hogy a termékek árváltozása egy általános monetáris tényezőtől, valamint az adott termékre jellemző specifikus piaci hatásoktól függ. Modern terminológiával élve a pénzügyi és a reálfolyamatok szétválasztása volt a cél.

Jevons gondolatmenete a következő volt: valamely i termék p_i árindexe $\lambda\varphi_i$ alakba írható, ahol λ a minden terméket érintő monetáris hatás, φ_i pedig a termékspecifikus komponens. Ennek alapján az infláció mértékét, λ -t a következő formulával mérhetjük:

$$\hat{\lambda} = \sqrt[n]{\prod_{i=1}^n \left(\frac{P_{i1}}{P_{i0}} \right)}.$$

E súlyozatlan formula tulajdonságai természetesen vitát váltottak ki, ennek egyik hangadója Edgeworth volt.

Fisher híres, 1927-ben megjelent könyvében („The making of index numbers”) bevezette az indexelméletbe a jól ismert atomisztikus vagy próbákön alapuló módszereket. Fisher fejtegetései nem emlíktetnek a sztochasztikus megközelítésre, azonban nála is felbukkan valami hasonló, amikor egy helyen azt mondja, hogy az infláció nem más, mint az ármozgásokban megmutatkozó közös trend.

De az is emlékeztet az SA-ra, amikor a bizonytalanságok kapcsán a „valószínű hibáról” írt egy másik helyen.

Diewert 1995-ben, a már említett tanulmányban, négy pontban kritizálta az SA-módszert. A szerzők részletesen foglalkoznak e négy szemponttal, amelyeket provokatívnak és konstruktívnak tartanak. A varianciára vonatkozó kritikára az a válaszuk, hogy a megfelelő elfogadható feltételek mellett azok feloldhatók. A súlyarányokat illetően azzal érvelnek, hogy azok meglehetősen állandók és így az idősorokat illetően nem okoznak problémát. Országok közötti összehasonlításban ugyan más a helyzet, de az is kezelhető a formula kismértékű átalakításával. A harmadik a probléma az „előregedés”. Ezzel részben együtt kell élni, de jelentősége a súlyok frissítésével csökkenthető. A súlyozást illető negyedik pont az előzők szerint lényegében megoldott.

A szerzők hangsúlyozzák, hogy az indexelméletben nincsenek tökéletes egyértelmű megoldások. Így a sztochasztikus formuláknak is vannak előnyeik és hátrányaik. Nagy előny az, hogy a relatív árak szórása lehetővé teszi a „bizonytalanság” becslését.

Theil 1967-ben írt tanulmányában azzal az esettel foglalkozott, amikor két, esetleg nagyon eltérő fogyasztási szerkezetű ország árszínterét összehasonlító „árindexét” keressük. Ilyen helyzetben nemcsak az egyedi árindexek szórása lehet nagy, hanem az egyedi volumenindexeké is. A megoldást egy harmadik, „semleges” súlyrendszer megválasztásában látja. A tanulmány a következőkben viszonylag terjedelmesen foglalkozik az országok közötti árszínter összehasonlítását leíró árindexekkel, amelyek problémái a fogyasztási szerkezet lényeges különbözőségeiből adódnak.

Mint tudjuk, az SA célja a „zajos” árakból következtetéseket levonni az infláció mértékére. A következőkben a tanulmány néhány további új eredményt ismertet az indexszámok

becsléséhez használt releváns információk természetéről és azok értékéről.

Tételezzük fel, hogy valamely x^* változó értékét előre akarjuk jelezni és van n torzítatlan előrejelzésünk, amelyek súlyozott átlaga X , akkor az optimális előrejelzés az lesz, amelynél az egyes becslések és az X négyzetes eltérése minimális. Ez az eljárás pedig a sztochasztikus indexhez vezet. Az eljárás részletei azt is jelzik, hogy minél kisebb az egyes előrejelzések szórása, annál jobb minőségű lesz az „előrejelzés”, ami az infláció mértékét jelzi. Hasonló eredményekre vezet, ha a mennyiségi pénzelméletből indulunk ki.

A Fisher-féle tényezőpróbát felhasználva, ha ismerjük például a fogyasztás értékének változását (M), akkor a volumenindexeket hasonlóan kezelve és felhasználva a jól ismert $P \times Q = M$ összefüggést, lehetőség van az árindex becslési hibájának csökkentésére.

Végül, a tanulmány két példával illusztrálja a sztochasztikus árindexek viselkedését. Mindkét példa ausztrál adatokat használ. Az elsőben öt termék (-csoport) 1952 és 1981 közötti adatait vizsgálja: becsüli az inflációt és a standard hibát. Ez alatt az idő alatt nem volt jelentős infláció Ausztráliában, a tanulmány az arányváltozásokra teszi a hangsúlyt, de bemutatja grafikusan a mintavételi hibát is. Negyedéves átlagban a lakással kapcsolatos költségek 4 százalékkal emelkedtek, míg a háztartási berendezések, tartós fogyasztási cikkek 3 százalékkal csökkentek az átlaghoz viszonyítva. Az élelmiszerek és egyéb cikkek ára az átlagnak megfelelően alakult.

A második példa jóval bonyolultabb. Azt mutatja be, hogyan lehet ezzel az indexszámítási módszerrel egy ország nemzetközi versenyképességét mérni. Tételezzük fel, hogy egy ország $c=1,2,\dots,n$ országgal van kapcsolatban és S_{ct} a nominális valutaárfolyam a t időpontban. Ha P_t és P_{ct}^* a hazai és a külföldi árszínter, akkor a valószínűségi árfolyam

$r_{ct} = p_t / s_{ct} P_t^*$. Ennek időbeni pozitív vagy negatív változása befolyásolja az adott ország versenyképességét. Ha a vásárlóerő-paritáson számított árfolyamváltozásokat az összes ország tekintetében átlagoljuk, akkor mód lesz az átlagos változáson belül elkülöníteni országcsoportokat, amelyekben viszonylag kedvezően, vagy kedvezőtlenül változott a versenyképesség, a „cserearány”. A szerzők Ausztráliának 26 országgal 1970 és 2002 közötti kereskedelmének adatai alapján mutatják be milyen mértékben növekedett Ausztrália versenyképessége. A példa arra is lehetőséget ad, hogy a szerzők rámutassanak, a konfidenciaintervallumok ismeretében miként lehet az elemzés részleteit finomítani.

Összefoglalásként a tanulmány megállapítja, az elmúlt évek során megnőtt a szakértők érdeklődése az indexszámok iránt. Ennek számos oka lehet: egyre nagyobb mennyiségben, minőségben, választékban álnak rendelkezésre javak, szolgáltatások és egyre több szempont szerint kell az árak változását mérni (core infláció, harmonizált árindex, nyugdíjasok fogyasztói kosara stb.). Az informatikai háttér is óriási fejlődésen ment át. A sztochasztikus indexeknek szerepük volt ebben a fejlődésben, „best fit” státust kaptak, és könnyen mérhetővé tették a bizonytalanság nagyságát. Mivel pedig ez a megközelítés is így vagy úgy a szokásos indexformulákhoz vezet, lehetővé vált azok új interpretációja.

A tanulmány szerzői meg vannak győződve arról, hogy az SA felhasználhatósága az itt felvázoltaknál sokkal szélesebb körű. Így például mérni lehet a teljes tényező termelékenységet (total factor productivity – TFP); felhasználhatók a nemzetközi összehasonlításokban; elemezhetők a pénzügyek az árak, az árfolyamok és a kamatok függvényében; elemezhetők a „strukturális” kamatok.

A tanulmány szükségképpen nem teljes körű, de befejezésül még négy kiegészítést

tesz. Elsőként azt említi, hogy az ár-, volumen- és értékindexek közötti összefüggésből adódó elemzések még nincsenek eléggé kihasználva. Másodikként a lakásárváltozások mérési módszereinek bővítését említik. Lehet ezekkel, a módszerekkel az ágazatok kibocsátásának relatív változásait is vizsgálni. Végül lehetne további kísérleteket folytatni a közgazdasági és sztochasztikus indexelmélet összevonására a hasznossági függvény egyes paramétereinek randomizálásával.

Marton Ádám

kandidátus, a Központi Statisztikai Hivatal ny.
osztályvezetője
E-mail: adam.marton@ksh.hu

Evertsson, M.:

Nemek szerinti reprodukció

(The reproduction of gender: housework and attitudes towards gender equality in the home among Swedish boys and girls.) – *The British Journal of Sociology*. 2006.évi 3. sz. 415–436. old.

A gyermekek szüleiket figyelve szerzik első benyomásait arról, hogy hogyan lehet, illetve kell a család életét megszervezni. A gyakran egyenlőtlen munkamegosztás folytán a gyermekek már viszonylag fiatal korukban különbséget tesznek a tipikusan „férfias”, illetve a „női jellegű” háztartási feladatok között.

A tanulmány a következő két kérdésre kívánt választ adni: 1. van-e eltérés a különböző nemű gyermekek által végzett házimunkák között, s ha igen, úgy követik-e a szülők munkamegosztásának mintáját; 2. befolyásolja-e a gyermekek magatartását szülőiknek a gyakorlatban tapasztalt viselkedése.

A téma korábbi vizsgálatai általában az Egyesült Államok adatain alapultak. (A szerző

gazdag bibliográfiára támaszkodva, igen részletesen ismerteti az eddigi szakmai álláspontokat.) A svéd tapasztalatok elemzése azért tarthat számot megkülönböztetett figyelemre, mert az ENSZ 1995-ben Svédországot nyilvánította annak az országnak, ahol a világon a legjobban érvényesül a nemek közötti egyenlőség. Ez ugyan még nem bizonyítja egyértelműen, hogy a családon belüli munkamegosztás Svédországban sokkal kiegyenlítettebb, mint bármely más országban, de tény, hogy a mintába bekerült férfiak és nők 80 százaléka a nemek családon belüli egyenlőségét olyan célnak nyilvánította, amelynek elérésére törek-szenek.

Svédországban 1968-ban végeztek először átfogó megfigyelést az életszínvonal vizsgálata érdekében. Ezt a panelfelvételt, amely a 18–75 éves korú népesség 1 ezrelékes mintáján alapult 1974., 1981., 1991. és 2000. évekre vonatkozóan megismételték. A legutóbbi, 2000. évi felvételt kiterjesztették a gyermekek és serdülőkorúak életszínvonalának speciális megfigyelésére is. Ebbe a mintába azok a 10–18 éves gyermekek (serdülők) kerülhettek be, akik havonta legalább tíz napot töltöttek az alapmintába tartozó felnőtt háztartásában. Ez 878 családból összesen 1304 gyermek részvételét jelentette (a válaszolási arány 85,3 százaléknak felelt meg). A megfigyeléssel közvetlen, személyileg nyújtott információkat kívántak nyerni mind a gyermekek, mind a szülők élethelyzetéről. E célnak megfelelően az elemzésbe csak a két, különböző nemű szülővel együtt élő fiatalokat vonták be. A minta elemszáma így 12 százalékkal csökkent (az egy szülővel, illetve azonos nemű szülőkkel együtt élő gyermekek figyelmen kívül hagyásával).

A gyermekek „beszélő kérdőívet” kaptak, s a kazettát meghallgatva válaszaikat előnyomatott szövegváltozatok útján rögzíthették a rendelkezésükre bocsátott füzetben. A felnőtt adatszolgáltatóknak postai úton eljuttatott kér-

dőívet kellett kitölteniük, s tőlük is kértek a gyermekekre vonatkozó információkat.

A szerző táblázatokba foglalva, részletesen bemutatja a regressziós számításoknál figyelembe vett független és függő változók nemenként is tagolt adatait, a számítások eredményeit és a főbb mutatók szórását. Néhány példa a – függő változóként kezelt – különböző házimunkák tartalmi jellemzésére. Elkülönítve vették számba a gyermekek által saját magukra fordított (saját szoba takarítása, saját ágy bevetése), illetve a család közös céljaira végzett tevékenységeit (többi szoba takarítása, főzés, asztalterítés, mosogatás, mosás, vasalás, viráglocsolás, kisebb testvérek gondozása, bevásárlás stb.). A család ellátására szolgáló házimunkákat a hagyományosan női feladatok egyik mutatójának tekintették. Jellemzésére összevont indexet képeztek, 1–8 között változó értékkel, vagyis minden fiú vagy leány 1 ponttal volt számításba vehető, ha hetenként legalább egyszer elvégezte az előbbieken felsorolt házimunkák valamelyikét. A házon kívüli munkákat – dummy változóként – vették figyelembe a hagyományosan férfi feladatok egyik jellemzőjeként. A szülők által végzett házimunka, illetve a szülőknek a családon belüli egyenlőséggel kapcsolatos magatartása a független változók körébe tartozott.

A vizsgálat eredményei szerint a 10–18 éves svéd gyermekek közül a legtöbben végeznek valamiféle házimunkát, de igen kevesen fordítanak erre számottevő időt. A fiúk és lányok átlagosan egyaránt egy-két óranyi házimunkát végeznek hetenként, s ez a munkamennyiség a korukkal együtt sem növekszik lényegesen. Ennek az lehet az egyik oka, hogy az idősebb tizenévesek lényegesen több időt töltenek otthonuktól távol, mint a kisebb gyermekek.

A különböző fajtájú háztartási tevékenységek tekintetében már jelentősek a nemek szerinti eltérések. A lányok sokkal több házon be-

lül, ellátási feladatot (főzés, mosogatás, takarítás) végeznek, mint a fiúk. Tulajdonképpen a házon kívüli munkák (javítások, szemétkihordás) tekinthetők az egyetlen olyan területnek, ahol a fiúk gyakrabban segítenek, mint a lányok. A háztartáson belüli, nemek szerint eltérő munkamegosztás tehát nemcsak a felnőttek körében szemmel látható, hanem a gyermekeknél is. Mindez arra mutat, hogy a hagyományosnak tekinthető magatartási mintákat a szülők – társadalmi hatásként – Svédországban is továbbadják gyermekeiknek.

A vizsgálat, a korábbi véleményekkel szemben, sok tekintetben árnyaltabb képet eredményezett. Nem bizonyult például megalapozottnak az a feltételezés, hogy szoros a kapcsolat a gyermekek és az azonos nemű szülőjük által végzett házimunka aránya között. Azokban a családokban ugyanis, ahol a házimunka nagy része az apára hárul, a lányok és a fiúk egyaránt többet segítenek. A szülők közötti egyenlőtlen munkamegosztás esetén mégis az a gyakoribb, hogy a lányok nagyobb arányban vesznek részt a háztartási feladatokban, mint a fiúk. Különösen jellemző ez a kisebb testvérek gondozására, ahol a lányok jelentős részét vállalhatják át az anyai feladatoknak.

Figyelemre méltó új megfigyelés viszont, hogy a szülők „tipikustól eltérő” magatartása meglepően erős hatást gyakorolhat a gyermekekre. Ennek egyik jellegzetes példája, hogy a lányok is gyakrabban vállalkoznak a házon kívüli munkákra, ha az anya az a személy a családban, aki a javítási, karbantartási munkákat végzi.

A fiúk beszámolóiból kitérünk, hogy azok a szülők, akik megkövetelik gyermekeiktől a bekapcsolódást a háztartási munkákba, többnyire a családon belüli egyenjogúságot is igen fontosnak tartják. A tapasztalatok szerint azonban ez a szülői felfogás nem befolyásolja döntően a fiúk magatartását. A lányok körében ilyen esetekben gyakrabban észlelhető pozitív

hatás. Általános érvényű megállapításnak bizonyult viszont, hogy minél iskolázottabb az anya, annál fontosabbnak ítélik mind a fiúk, mind a lányok a családon belüli egyenlőség érvényesülését.

A tapasztalatokat összegezve megállapítható, hogy – bár csak gyenge kapcsolat mutatható ki a gyermek és a vele azonos nemű szülő tevékenysége között – a svéd gyermekek körében világosan észlelhető a házimunkavégzés nemek szerinti elkülönülése. A lányok, a felnőtt nőkhöz hasonlóan, felelősebben reagálnak a házimunkával járó kötelezettségekre, és a terhek nehezülése esetén ők vállalnak jelentősebb részt azok elvégzéséből. A vizsgálat arra nem tudott választ adni, hogy a fiúk csekélyebb részvállalása személyes vonakodásuknak tulajdonítható-e, vagy abból ered, hogy a családon belüli munkamegosztás kialakításakor a szülők eleve az ő javukra döntenek. Ha a szülők felelősségteljesen járnak el a gyermekek házimunkára nevelésében, a fiúk is átérzik a nemek családon belüli egyenlőségének fontosságát. Arra a kérdésre azonban, hogy ez valóban kifejezésre jut-e a magatartásukban és felnőttkorukban is érvényre jut-e, további vizsgálatoknak kell választ adniuk.

Tűz Lászlóné,

a Központi Statisztikai Hivatal ny. osztályvezetője

A férfiak és a nők számítógép- és internethasználatát jellemező különbségek

(Gender differences in the use of computers and the Internet.) – Statistics in focus. 2007. évi 119. sz.

A kiadvány a 2006. évi közösségi ICT-felvétel eredményei alapján a számítógéphasználat nem specifikus vonásait összegzi. Ezzel kapcsolatosan a főbb megállapítások a következők.

A számítógépet napi gyakorisággal használó nők és férfiak aránya a fiatal korosztályban közel hasonló volt 2006-ban (62, illetve 67 százalék), viszont jelentősen különbözött a 25–54 és az 55–74 évesek körében.

A napi internethasználat a fiatal férfiakra egy árnyalatnyit jobban jellemző, mint a fiatal nőkre (53, illetve 48 százalék). Az idősebb generációkból kevesen interneteznek, viszont a különbség a férfiak javára jelentős. (Az 55–74 éves férfiak 18 százaléka, a nők 9 százaléka volt napi internethasználó.)

A közepes, illetve magas szintű számítógépes ismerettel rendelkező nők aránya valamennyi korcsoportban alacsonyabb volt, mint a férfiaké. A nők között a számítógépes szakemberek aránya igen alacsony (0,7%), és 2001 és 2006 között nem változott, miközben ez az arány a férfiak esetében 2,3-ról 2,6 százalékra nőtt.

Az EU lakosságának igen jelentős része ma rendszeres számítógép-használó. A 2006. évi az ICT használatára irányuló lakossági felvétel szerint a 16–74 éves férfiak 48 százaléka, a nők 39 százaléka használ legalább naponta egyszer számítógépet. Az egyetlen ország ahol a nők a nagyobb arányú számítógép-használók Ciprus, Bulgária, Észtország és Litvánia esetében nincs e téren érdemi különbség a két nem között, míg a többi tagországban a férfiak előnye jelentős, a legnagyobb Luxemburgban, ahol eléri a 24 százalékpontot. Az új tagországokban a nők és férfiak számítógép-használati szokásaiban kisebb az eltérés, mint a régi tagországokban. A napi gyakorisággal számítógép-használók aránya 60 százalék, vagy e feletti Dániában, Hollandiában, Svédországban, Izlandon, és Norvégiában, míg a másik pólust Görögország jelenti 20 százalékos értékével.

A napinál ritkább gyakorisággal számítógépet-használók aránya a férfiak és nők esetében közel hasonló. A számítógép-használatban

a nők és férfiak közötti különbség az életkor emelkedésével párhuzamosan nő. A 16–24 éves fiúk között a hetente legalább egy alkalommal számítógépet használók aránya csak 2 százalékkal magasabb, mint a hasonló korú lányoknál. A napi használat a korcsoportban 75 és 40 százalék között változik. A legmagasabb érték Svédországot, a legalacsonyabb Bulgáriát jellemzi. A 25–54 évesek kisebb arányban használnak számítógépet, mint a fiatalok. A nők esetében a napi használók aránya 47 százalék a férfiaknál pedig 54 százalék. Nyolc tagországban (valamennyi új tagország) a nők inkább használják a számítógépet mint a férfiak, míg néhány tagországban (például Ausztria, Egyesült Királyság, Svédország) a vizsgált korosztályban a férfiak előnye jelentős, meghaladja a 10 százalékpontot. Az idősekre a számítógép napi használata kevésbé jellemző és jelentős a férfiak és nők közötti különbség. Összességében az 55–74 éves férfiak 26 százaléka (Dániában 54, Bulgáriában 5 százalék), a nők 15 százaléka (Dániában 45, Görögországban 2 százalék) használta csak naponta a számítógépet.

Az internethasználatban a férfiak és nők közötti különbség mértéke a számítógép-használatáéhoz hasonló, azzal hogy a férfiak e téren minden tagországban előnyben vannak. A nemenkénti különbség az új tagországokban az átlagosnál kisebb, de az internetezés is kevésbé elterjedt. Az internethasználatban az északi országok előnye jelentős, valamint az életkor növekedésével növekszik a nők és férfiak közötti különbség.

2006-ban a férfiak fele értékelte számítógépes tudását magas szintűre, szemben a hasonló korú nők 30 százalékaival. Az előbbi arány hét tagországban meghaladta a 60 százaléket, 6 országban pedig az 50 százaléket. Ezzel szemben a lányoknál csak Ausztriában és Szlovéniában volt 50 százalék feletti érték. A 25–54 évesek számítógépes ismerete jóval el-

marad a fiatalokétól, itt csak minden harmadik férfi és a nők 18 százaléka tekintette magát magas szintű számítógépes ismeretekkel rendelkezőknek. A nők és a férfiak közötti különbség hat tagországban a 20 százalékpontot is meghaladta. Az 55–74 éves nők 3 százaléka a férfiak 12 százaléka minősítette informatikai tudását magas szintűnek.

Az Unióban növekszik a számítástechnikai munkát (ISCO kód 213 és 312) végzők száma. Ők 2006-ban már az összes férfi foglalkoztatott 2,6 százalékát jelentették. Az informatikusi munkakör férfi dominanciáját mutatja, hogy ugyanakkor a nőknek mindössze 0,7 százaléka dolgozott ilyen területen. A nemenkénti szegre-

gáció Bulgária, Hollandia, Ausztria és Portugália estében a legjelentősebb (körülbelül 5:1), míg Románia és Lettország képviseli a másik pólust 2:1 arányával. A férfiak és nők között e téren fennálló különbség 2001 és 2006 között még nőtt is. Mivel a 40 év alatti nők alig valamivel nagyobb hányada dolgozik informatikusként, mint a 40 felettieknek (0,8, illetve 0,5 százalék) miközben a férfiaknál a generációs eltérés jelentős (3,5, illetve 1,8 százalék), ez a tendencia tartósan fenn fog maradni.

Lakatos Judit

PhD, a KSH főosztályvezetője
E-mail: judit.lakatos@ksh.hu

Kiadók ajánlata

DASGUPTA, A. [2008]: *Asymptotic theory of statistics and probability*. (A statisztika és valószínűség aszimptotikus elmélete.) Springer. New York.

A könyv a klasszikus és a modern nagymintás elmélet enciklopédikus feldolgozása, mely érthető stílusban mind statisztikai, mind valószínűségi kérdésekkel és eszközökkel foglalkozik. Nagy hangsúlyt fektet a problémák elméleti hátterének megvitatására és a teorémák hatásának és fontosságának megvilágítására. A könyv 34 fejezete a témák széles skáláját öleli fel, majd 600 oktatásra és gyakorlásra szánt feladatot, valamint 300 kidolgozott példát is tartalmaz. Megtalálható benne továbbá egy hatalmas gyűjtemény, mely 300 hasznos egyenlőtlenséget tartalmaz a valószínűség-számítás és a lineáris algebra témaköréből, továbbá számos forrásból összegyűjtött elemzéseket tartalmaz, amelyek értékes segítséget jelentenek statisztikával, matematikával és valószínűség-számítással foglalkozó szakemberek számára.

Oktatási anyagként éppúgy használható, mint kutatói szakirodalomként, vagy önálló olvasmányként a klasszikus nagymintás elmélet területeiről, valamint kitekintést nyújt a modern témák legújabb eredményeire. A könyv egyedülálló olyan alapvető témák részletes ismertetésében, mint a centrális határeloszlás tételek számos alkalmazása, likelihood-alapú módszerek, az illeszkedés jósága, magasabb rendű asszimptotikus közelítések, éppúgy, mint a legmodernebb témák, mint a bootstrap, a függő adatok, a bayesi asszimptotikák, a nemparametrikus sűrűségbecslés, a vegyes modellek és többszörös tesztelés, továbbá az első- és másodfajú hibákkal kapcsolatos ismeretek. A könyv irodalomjegyzéke bőséges szakirodalmat tartalmaz az összes témára, valamint a legújabb publikációkat.

DUDOIT, S. – VAN DER LAAN, M. J. [2008]: *Multiple testing procedures with applications to genomics*. (Többszörös teszteljárások genomtani felhasználásokkal.) Springer. New York.

A könyv a többszörös hipotézisvizsgálat általános módszertanának elméleti alapjait fekteti le, és megvitatja ezek szoftveres beilleszthetőségét az R- és SAS környezetbe. A módszereket az orvosbiológiai és genomkutatások egy sor tesztelési problémájára használják, beleértve az expresszált és koexpresszált gének azonosítását az intenzív teljesítményű gén-expressziós kísérletekben, mint a mikrorendszerű kísérletek, asszociációs tesztek a gén-expressziós mérések és a biológiai annotációs metaadatok (például génontológia), szekvenencia-elemzés és a komplex jellemzők genetikai feltérképezése egyedi nukleotidok közötti polimorfizmus használatával.

A munka egyaránt célozza azokat a statisztikusokat, akik érdeklődnek a többszörös teszt-elmélet iránt és azokat, akik többdimenziós tesztproblémákkal foglalkoznak.

A könyv másodlagos mintavételi eljárásokon (resampling) alapuló egylépcsős és többlépcsős tesztmódszereket javasol az elsőfajú hibák széles körének ellenőrzésére, melyeket elvárt értékek meghatározásaként az elsőfajú hibák és a visszautasított hipotézisek tetszőleges függvényeinek kritikus tartományaként és várható értékeiként értelmez (például hamis találati arány). A létező megközelítésektől eltérően az eljárás a kapcsolt nulla valószínűség-eloszlású tesztstatisztikákon alapul és az elsőfajú hiba ellenőrzését biztosítja a tesztproblémában beleértve az általános adatgeneráló eloszlásokat (változók közötti tetszőleges függő struktúrákkal), a nullhipotézist és a tesztstatisztikákat. A többszörös teszteredményeket az elutasított tartományokban, a paraméter konfidencia-tartományában és a korrigált p -értékekben adják meg.

KOSOROK, M. R. [2008]: *Introduction to empirical processes and semiparametric inference*. (Bevezetés a tapasztalati folyama-

tokba és a félparaméteres következtetésekbe.) Springer. New York.

A könyv teljes körű és egységes bevezetést tartalmaz a tapasztalati folyamatokhoz és a félparaméteres következtetésekbe. Ezek a hatásos kutatói technikák meglepően hasznosak a komplex modellek alapján történő statisztikai levezetésének módszeréhez, és ezen módszerek megértéséhez. A könyv célközönsége statisztikusok, biostatistikusok és olyan matematikai statisztikai alapokkal rendelkező kutatók, akik érdeklődnek a tapasztalati folyamatok és a félparaméteres következtetések iránt, és könnyen érthető, fokozatos bevezetést keresnek a témához. A könyv kutatói szakirodalomként és tankönyvként egyaránt használható.

A munka három részből áll. Az első könyvben foglalt elméletek áttekintését tartalmazza minimális technikai részletekkel. A második és a harmadik rész tárgyalja teljes mélységében a két fő témát: a félparaméteres következtetést és a gyakorlati folyamatokat. A két téma összefüggését a szerzők is teljes mértékben bemutatják és hangsúlyozzák. Minden rész zárófejezete esettanulmányokat tartalmaz, konkrét példákkal melyek alátámasztják az elméletek fejlődését. Az utolsó két rész mindegyike tartalmaz egy olyan fejezetet, amely bemutatja a szükséges matematikai előzményeket. Ennek során nem a technikai részletekre helyezik a hangsúlyt és minden fontosabb fogalmat teljes példákkal támasztanak alá. A minden fejezet végén megtalálható házi feladatok segítik az olvasót, hogy mélyebb betekintést nyerjen a problémákba

O'QUIGLEY, J. [2008]: *Proportional hazards regression*. (Arányos hazard regresszió.) Springer. New York.

A túléléselemzés témakörében az arányos hazard modellek és általánosításai már olyan nagy mértékben jelennek meg, hogy nem lehet a témáról kurzust tartani anélkül, hogy leg-

alább a kurzus felét nem ennek a területnek szentelik. A könyv a modern túléléselemzés alapjául szolgáló nem arányos kármodell és az azok speciális esetét képező arányos kármodell széles körének elméletére és alkalmazására összpontosít. Más ezzel a témakörrel foglalkozó könyvektől eltérően a könyvben a hangsúly nem a sztochasztikus integrálok és martingálok elméletének bemutatására esik. Helyettük, mivel a levezetések számítási eljárásokon alapulnak és a martingálok elméletét fedik le, sokkal nagyobb hangsúlyt fektetnek olyan hagyományos eredményekre, mint például a centrális határeloszlás tétel. Ez a hangsúlyeltolódás teszi lehetővé, hogy a könyv jóval több figyelmet szenteljen a modellezés gyakorlatának. A különböző modellek felhasználása, gyakorlati interpretációja, bármely modell előrejelző képessége, modellalkotás és modellválasztás éppúgy, mint a rosszul specifikált modellek problémája kiemelt figyelmet kapnak. A könyv azokat célozza, akiket mind az elmélet mind a felhasználási gyakorlat érdekel. A mű sok példát és illusztrációt tartalmaz. A szükséges matematikai és statisztikai alapokat biztosítja azok számára is, akik számára új a terület, így az anyag különböző szintű tudással rendelkezők számára is feldolgozható.

Duchateau, L. – Janssen, P. [2008]: *The frailty model*. (A frailty modell.) Springer. New York.

Klaszterezett túlélési adatokkal sokféle tudományterület foglalkozik beleértve az orvos- és állatorvos-tudományt, a biológiát, a epidemiológiát, a közegészségügyet és a demográfiát. A frailty modellek erőteljes eszközök a klaszterezett túlélésadatok elemzéséhez. A témáról megjelent nagyszámú kutatói publikáció ellenére viszonylag kevés statisztikai szoftvercsomag tartalmaz frailty modelleket.

A létező szakirodalomból gyakorló statisztikusok és végzős diákok számára nehéz a frailty modellek jó megismerése. A könyv mélységeiben tárgyalja és értelmezi a frailty modellek módszertanát az olvasók számára. Ismerteti a parametrikus és a félpáraméteres frailty modelleket és az AFT-modelleket. Bemutatja a frailty modellekhez kapcsolódó gyakorlati módszereket, mint az EM algoritmus, büntetett likelihood technikák, Laplace integrál és a bayesi technikák. A munka a hierarchikus adatokra alkalmazható és magában foglalja a frailty modelleket is. A szerzők valós életből vett példákkal mutatják be a frailty modellek különös illeszthetőségére és az eredmények értelmezhetőségére vonatkozó eljárásokat. A könyvben megtalálható kidolgozott példához kapcsolódó eljárások letölthetők a kiadó weboldaljáról, melyek többségét az R és a Winbugs ingyenes programcsomagokhoz fejlesztették ki. A könyv a klasszikus túléléselemzés néhány alapvető elméletének áttekintésével kezdődik, összegyűjtve mindazt, ami szükséges a bonyolultabb frailty modellek tanulmányozásához.

Társfolyóiratok



A FRANCIA GAZDASÁGI ÉS PÉNZÜGYMI-
NISZTÉRIUM ÉS A STATISZTIKAI ÉS
GAZDASÁGKUTATÓ INTÉZET FOLYÓIRATA

2007. ÉVI 405-406. SZÁM

Flacher, D. – Pelletan, J.: Az ipar fogalma és mérhetősége. Eredet, korlátok és lehetőségek. Alkalmazás az ipar változásának vizsgálatához.

Leclair, M. – Roux, S.: Relatív termelékenység és a rövid távú munkaviszony alkalmazása cégeknél.

Gonzalez, L. – Picart, C.: Diverzifikáció, újragondolás és a segítségnyújtás szintje csoportokon belül (1993–2000).

Colombier, N. – Masclat, D.: A családi környezet fontossága mint az önfoglalkoztatás meghatározója.

Nefussi, B.: A külkereskedelem térhódítása a vegyiparban.

Thévenot, C.: A K+F hálózat nemzetköziesedése, a vállalati kapcsolatok megközelítése.



AZ EGYESÜLT ÁLLAMOK
MATEMATIKAI STATISZTIKAI INTÉZETÉNEK
FOLYÓIRATA

2007. ÉVI 3. SZÁM

Bédard, M. – Fraser, D. A. S. – Wong, A.: Nagyobb pontosság elérése bayesi- és gyakorlati becsléseknél: nagymintavétel-elmélet kismintavételes valószínűséghez.

Bayarri, M. J. – Castellanos, M. E.: Bayes-i ellenőrzés a hierarchikus modellek második szintjén.

Friendly, M.: A.-M. Guerry Franciaország morális statisztikája c. műve: Kihívások a többváltozós területi elemzések terén.

George, E. I.: Ingram Olkin emlékezete.

Becker, B. J.: Többváltozós metaelemzések: Kiegészítések Ingram Olkin eredményeihez.

Arnold, B. C.: Majorizáció: itt, ott és mindenhol.

Becker, B. J. – Wu, M.-J.: Regressziós görbék metaelemzésekben.

Drton, M. – Perlman, M. D.: Többszörös ellenőrzés és hibakeresés a Gauss-féle gráfok modellválasztásában.

Sampson, A. R.: Beszélgetés Ingram Olkinnal.



A CSEH STATISZTIKAI HIVATAL
FOLYÓIRATA

2008. ÉVI 1. SZÁM

Dyba, K.: Átmenet, bruttó hazai termék és gazdasági jólét: Csehország 1989–1995.

Jilek, J. – Vojta, M.: Az össztermelékenység szerepe a cseh gazdaságban.

Klas, A.: Oktatás, kutatás és fejlesztés: új tényezők a gazdasági fejlődésben a második világháború után.

Novák, J.: A szabadfoglalkozásúak társadalmi-gazdasági státusza. (3. rész.)

Trešl, J. – Blatná, D.: Válogatott európai részvénytatók statisztikai elemzése.

Češka, J.: A Cseh-Morva Protektorátus statisztikai áttekintése.

Drápal, S.: Együttes UNECE/Eurostat találkozó a népesség- és lakásösszeírásról. Csehország gazdasági fejlődési mutatói.



AZ OROSZ ÁLLAMI STATISZTIKAI
BIZOTTSÁG FOLYÓIRATA

2007. ÉVI 11. SZÁM

Dolgoplov, P. I. – Konratyeva, I. I.: „Az orosz állami statisztikai elszámolásról és az állami statisztika rendszeréről” szóló föderációs törvény vázlatáról.

Kotlyarevskaya, T. I.: A statisztikai tevékenység perspektívájának és jelenlegi tervezésének kérdései.

Petrikov, A. V.: Az „Agrár és ipari komplexum fejlesztése” c. prioritásos országos projekt monitoringjának fő eredményei.

Kozlov, M. P.: Kisvállalkozások az „Agrár és ipari komplexum fejlesztése” c. prioritásos országos projekt megvalósításának feltételei mellett.

Barsukova, S. Yu.: Az „Agrár és ipari komplexum fejlesztése” egy ötlet és megvalósítás.

Bokun, N. C.: A kis- és középvállalkozások tevékenységével kapcsolatos eredmények mérési problémái.

Chistenko, E. E.: Rejtett infláció: a kategória kezelése és becslési módszerek.

Bushueva, L. I.: Modern statisztikai beszámolók az innovációról és javításának lehetőségeiről.

Yusupova, G. F.: Az árak normalizálása a régiókban: a régiók közötti kereskedelem és a szállítási tarifák szerepe.

Kudryavtseva, R. M. – Starikova, O. S.: Bruttó regionális termék, mint gazdasági lehetőség szociális programok megvalósítása esetén.

Sin'kova, L. P.: A Khakasia Köztársaság társadalmi fejlettségi és biztonsági szintjének becslése.

Kadiatu, D. J. S.: A nők gazdasági tevékenységének statisztikai jellemzői a Mali Köztársaságban.

Gasanly, Y. – Suleimanov, N.: A hozzáadottérték-adó statisztikai elemzése és előrejelzése az Azeri Köztársaságban.

Yuzbashev, M. M. – Mihailova, T. M.: Rendszerhatás a gazdaságban és az életben.

2007. ÉVI 12. SZÁM

Pashintseva, N. I.: Az orosz központi államigazgatás és az önkormányzatok ellátása statisztikai információval.

Ivanov, Y. N.: Regionális elszámolások a FÁK országaiban.

Kandilov, V. P.: Humánpotenciál, mind a terület erőforrás-potenciáljának egyik tényezője.

Stepanov, S. V.: Hibaköltségek és a felvételi minőségjavítás kiadásai.

Beninel', F. – Grun-Reom, M.: Objektumok azonosításának értékelési szabályaival összefüggő hibák költségei.

Sablina, E. A.: Az orosz pénzügyi rendszer modern státuszának statisztikai elemzése.

Oktyabrskiy, P. Y.: Oroszország ma: jövedelmünk.

Plyshevskiy, B. P.: A GDP dinamikája és szerkezete.

Kiselnikov, A. A.: A statisztikai információ szerepe egy szibériai agro-ipari komplexum utólagos fejlesztési stratégiájának kidolgozásához.

Emelyanov, V. V.: A 2006-os oroszországi mezőgazdasági census felhasználása a déli föderációs körzet közigazgatási tevékenységeihez.

Sviridova, N. V.: A csoportosítások használatának módszertani kérdései a különböző tevékenységű szervezetek pénzügyi feltételeinek összehasonlító elemzésénél.

Kholiser, N. V.: A költségplusz-módszer használata az immateriális javak becslésénél.