

Statisztikai Szemle

A KÖZPONTI STATISZTIKAI HIVATAL
TUDOMÁNYOS FOLYÓIRATA

SZERKESZTŐBIZOTTSÁG:

DR. BELYÓ PÁL, ÉLTETŐ ÖDÖN, DR. HARCSA ISTVÁN, DR. HUNYADI LÁSZLÓ (főszerkesztő),
DR. JÓZAN PÉTER, DR. MÁTYÁS LÁSZLÓ, NYITRAI FERENCNÉ DR., DR. OBLATH GÁBOR,
DR. PUKLI PÉTER (a Szerkesztőbizottság elnöke), DR. RAPPAI GÁBOR, DR. SIPOS BÉLA,
DR. SPÉDER ZSOLT, DR. SZÉP KATALIN, DR. SZILÁGYI GYÖRGY, DR. VITA LÁSZLÓ

84. ÉVFOLYAM 3. SZÁM

2006. MÁRCIUS

*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. Hunyadi László
Osztályvezető: Dobokayné Szabó Orsolya
Kiadja: a Központi Statisztikai Hivatal
A kiadásért felel: dr. Pukli Péter
4421 – Akadémiai Nyomda
Martonvásár, 2006
Felelős vezető: Reisenleitner Lajos

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

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

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

Előfizetésben terjeszti a Magyar Posta Rt. 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., Keleti Károly u. 10. Telefon: 212-4348

Tartalom

Tanulmányok

A sokasági értékösszeg becslése a könyvvizsgálatban – <i>Lolbert Tamás</i>	225
Közlekedési infrastruktúra és jólét Kelet-Magyar- országon. – <i>Ohnsorge-Szabó László</i>	249
Az anyagáramlás-elemzés (statisztikai) módszertani kér- dései I. – <i>Szabó Elemér – Pomázi István</i>	271

Műhely

Nemlineáris függvények illesztésének néhány kérdése – <i>Tóth Zoltán</i>	284
---	-----

Fórum

Hírek, események	303
------------------------	-----

Szakirodalom

Könyvszemle

David Feedman – Robert Pisani – Roger Purves: Sta- tisztika – <i>Hunyadi László</i>	306
--	-----

Folyóiratszemle

Fairlie, R. W.: Vállalkozáselemzés az 1979. évi lon- gitudinális felvétel alapján az Egyesült Államok- ban – (Nádudvari Zoltán)	311
Bellmann, L. – Hilpert, M. – Kistler, E. – Wahse, J.: A demográfiai átalakulás hatása a munkaerőpiac- ra és a vállalatokra – (Waffenschmidt Jánosné) ..	313
Bainster, J.: feldolgozóipari foglalkoztatás Kínában – (Dévai Péter)	315
Kiadók ajánlata	318
Társfolyóiratok	320

A sokasági értékösszeg becslése a könyvvizsgálatban

Lolbert Tamás,
az Állami Számvevőszék szám-
vevője, a Budapesti Corvinus
Egyetem PhD-hallgatója
E-mail: lolbertt@asz.hu

A tanulmány célja, hogy áttekintést nyújtson az értékösszeg becslésére használt módszerekről, különös tekintettel a könyvvizsgálatban alkalmazottakra. A könyvvizsgálat azért tekinthető különleges területnek, mivel az általánosan használt, első és második momentumokra alapozott becslésekhez „jól viselkedő” eloszlásokra van szükség, és a pénzügyi beszámolókból található hibák eloszlása az empirikus vizsgálatok szerint nem ilyen. A terület specialitásainak ismertetését követően a tanulmány bemutatja a leginkább elterjedt becslési eljárásokat.

TÁRGYSZÓ:
Mintavétel.
Nem egyenlő valószínűséggel történő kiválasztás.
Pénzügyi alkalmazások, pénz- és értékpia.

A könyvvizsgálatban a leginkább tipikusnak nevezhető feladat az, hogy egy pénzügyi kimutatásról el kell dönteni, tartalmaz-e *lényeges hibát* (material error). Egy hiba lényegessége (materiality) a klasszikus értelmezés szerint abból fakad, hogy miatta már érdemben módosulnak a pénzügyi kimutatás alapján hozott döntések. A jelen tanulmány, és a legfontosabb gyakorlati alkalmazások szempontjából a lényeges hiba mindig a főösszeg (vagy fontosabb részösszegek) kimutatott és valós értékének¹ egy tolerálható mértéket meghaladó eltérését jelenti.²

Mivel a könyvvizsgálat meglehetősen távol áll a statisztika szokásos alkalmazásaitól, célszerűnek látszik egy kicsit részletesebben foglalkozni az alapfogalmakkal. Lényegesnek tekinthetjük a hibát például, ha az eltérés meghaladja az 1 millió forintot, vagy lényeges a hiba, ha az eltérés meghaladja a főösszeg 2 százalékát. Ezt a kritikus összeget (vagy százalékot) más néven *lényegességi küszöbnek* is hívják. Mivel semmiféle megszorítást nem jelent, a továbbiakban a lényegességi küszöb mindig a főösszegre vonatkozik, és a főösszeg százalékában (tehát nem abszolút összegben) van megadva.

Ez a definíció a következő példával érzékeltethető legkönnyebben. Tekintsük a számviteli törvényben leírt „A” típusú mérleg egy egyszerűsített formáját:

Eszközök (aktívák)	Források (passzívák)
A. Befektetett eszközök	D. Saját tőke
I. Immateriális javak	I. Jegyzett tőke
II. Tárgyi eszközök	...
III. Befektetett pénzügyi eszközök	VII. Mérleg szerinti eredmény
B. Forgóeszközök	E. Céltartalékok
I. Készletek	F. Kötelezettségek
II. Követelések	I. Hátrasorolt kötelezettségek
III. Értékpapírok	II. Hosszú lejáratú kötelezettségek
IV. Pénzeszközök	III. Rövid lejáratú kötelezettségek
C. Aktív időbeli elhatárolások	G. Passzív időbeli elhatárolások
Eszközök összesen	Források összesen

¹ A „valós érték” kifejezést a tanulmány nem a számviteli törvényben (2000. évi C. Tv 3.§. (9) 12.) meghatározott értelemben használja. A továbbiakban „valós érték” alatt az auditor által végzett *teljes körű* ellenőrzés után kapott „helyes” értéket kell érteni. Természetesen *ez az érték hipotetikus*, hiszen teljes körű ellenőrzésre nem kerül sor.

² A lényegesség az itt leírtnál jóval összetettebb fogalom, azonban további dimenzióit alapvetően nem a statisztika eszközeivel szokás megragadni.

Tegyük fel, hogy a mérlegben szereplő eszközök, mint kimutatás megbízhatóságáról kell véleményt nyilvánítani. Tegyük fel továbbá a példa kedvéért, hogy lényegesnek tekinthető az Eszközök összértékének 2 százalékos, a Befektetett eszközök értékének 1 százalékos, a Forgóeszközök értékének 5 százalékos, az Aktív időbeli elhatárolások értékének 2 százalékos, és végül a Készletek és a Követelések együttes értékének 10 százalékos eltérése. Az egyes lényegességi küszöbököt külön-külön kell vizsgálni. Könnyen látható, hogy a Befektetett eszközök, az Aktív időbeli elhatárolások, illetve a Készletek és a Követelések együttes értékének vizsgálata „tisza eset”, hiszen nem tartalmaznak további, lényegességi küszöbököt tartalmazó bontásokat. Ezzel szemben a Forgóeszközök, és az Eszközök összesen értékelése ebben az értelemben többlépcsős folyamat. A tanulmány elején csak a tiszta esetekkel foglalkozunk, és csak később térünk ki röviden az összetett esetek kezelési módjára.

A tanulmány célja azon klasszikus (Neyman–Pearson-elvet követő) statisztikai eszközök ismertetése, melyeket az elmúlt 25 évben fejlesztettek ki és jelenleg is használnak a pénzügyi beszámolók megbízhatóságának megítéléséhez. A problémát a statisztika nyelvére lefordítva a következő feladattal állunk szemben: minta alapján becsülni kell a sokasági értékösszeget, és ezt össze kell vetni a kimutatásban szereplő összeggel.

A következőkben tehát tekintsük az Y_i és X_i ($i=1 \dots N$) páros sokaságot, ahol Y_i jelöli az N tételből álló kimutatásban szereplő értékeket, és X_i ezek valós, de nem ismert értékét. Ez alapján a könyvvizsgáló feladata annak eldöntése egy előre adott (például 95 százalékos) bizonyossággal, hogy a teljes könyv szerinti érték ($Y = \sum_i Y_i$) és a valós érték ($X = \sum_i X_i$) különbsége hogyan viszonyul a lényegességi küszöbhez. Amennyiben az eltérés nem haladja meg a lényegességi küszöböt, elfogadja a kimutatást, ellenkező esetben elutasítja.³

A jóhiszemű feltevés szerint a könyvvizsgáló minden általa megvizsgált Y_i esetén képes X_i pontos megadására, de mivel megelégszik a részleges bizonyossággal, ezért döntését az összesen N tételből n megvizsgálásával fogja meghozni. A mintába kerülő n tételt – az általános sokasági tételektől megkülönböztetendő – Y_i és X_i kisbetűs változataival (y_i és x_i) jelöljük. Adott tétel könyv szerinti és a valós értéke segítségével számíthatjuk a következő két mutatót:

1. $d_i = y_i - x_i$ ($D_i = Y_i - X_i$), a minta (sokaság) i -edik elemében levő hiba vagy eltérés (error vagy deviation), ami a minta esetében

³ A könyvvizsgálat nem csak elfogadó és elutasító véleménnyel végződhet. Ahogyan azt a lényegesség definíciójával kapcsolatos 2. számú lábjegyzet is tartalmazza, a könyvvizsgálat a mintavételi módszereken kívül sok egyéb eljárást is használ, melyek esetleg feleslegessé is tehetik a mintavételt, továbbá a pénzügyi kimutatás egyes részeire adott lényegességi küszöbök aggregálásával sok esetben nem adható sem egyértelmű elfogadó, sem egyértelmű elutasító vélemény. A mi szempontunkból azonban, figyelembe véve a lényegesség általunk használt leegyszerűsített definícióját, megengedhető az ilyen egyszerűsítés.

ismert, a sokaság általános elemére pedig nem ismert, de létező érték;

$$2. t_i = \frac{y_i - x_i}{y_i} = \frac{d_i}{y_i} \quad (T_i = \frac{Y_i - X_i}{Y_i} = \frac{D_i}{Y_i}), \text{ a minta (sokaság) } i\text{-edik}$$

elemének szennyezettsége, tehát a könyv szerinti értékhez viszonyított relatív hibája (tainting).

Ha jól megfigyeljük, azonnal kitűnik a leírt modell legnagyobb hibája: ez a módszer nem alkalmas a „kifelejtett” tételek felderítésére. A továbbiakban tehát feltesszük, hogy nincsenek ilyen, „kifelejtett” tételek⁴ és csupán a tételek értékelése lehet hibás.

1. A könyvvizsgálatban előforduló populációk főbb statisztikai jellemzői

Ahhoz, hogy megértsük, miért is problematikus a könyvvizsgálatban az értékösszeg becslése, mindenképpen be kell mutatni a sokaság (populáció) jellegzetességeit. Ezzel kapcsolatosan az 1960-as, az 1970-es és az 1980-as években sok tanulmány született, melyek fő eredményeit ez a fejezet foglalja össze.

Az első szembetűnő jelenség, hogy a tételek túlnyomó része helyes, azaz nem tartalmaz hibát. Ez azzal jár, hogy a megvizsgált mintának csak minimális része tartalmaz érdemi információt a hibákról. *Johnson–Leitch–Neter* [1981] tanulmányából kiderül, hogy az általuk vizsgált adatállományokban a „Vevők” tételek (a B/II. „Követelések” egy alcsoportja) hibaarányának mediánja 0,024 (a kvartilisek $Q_1=0,004$ és $Q_3=0,089$), míg a „Készletek” ellenőrzésekor ugyanezek a mutatók $Q_1=0,073$, $Q_2=0,154$ és $Q_3=0,399$. Ezért például a nagy számok törvénye szerint a vevőállományból vett 500 elemű (tehát nagy) mintánál körülbelül 12 darab tétel információ-tartalma alapján kell az egész sokaságról nyilatkozni. A tanulmány szerint a nem nulla hibák (eltérések) eloszlása lényegesen eltér az egyes beszámoló-területeken. Míg a vevők esetén például szinte kizárólag csak túlértékelések (overstatement) szerepelnek, addig a készleteknél az alul- és túlértékelések körülbelül fele-fele arányban fordultak elő. Más tanulmány (*Ham–Lassel–Smieliauskas* [1985]) kitért a „Szállítók” tételeire is, ahol az alulértékelés (understatement) volt a tipikus.

A második specialitása ezeknek a sokaságoknak, hogy a nagyobb könyv szerinti értékű tételek nagyobb valószínűséggel tartalmaznak hibát, ám a relatív hiba (elté-

⁴ Ezt azért tehetjük fel, mert a könyvvizsgálat során a könyvvizsgáló egyéb módon már megbizonyosodott arról, hogy a szervezet belső eljárásai garantálják-e a kimutatások teljes körűségét.

rés/könyv szerinti érték) nagysága nincs szignifikáns kapcsolatban a könyv szerinti értékkel. Ezen felül, az eltérés szórása a könyv szerinti értékkel növekszik. Amennyiben az adott tétel egyes pénzegységeit, pontosabban a tétel ezekhez rendelt relatív hibáját tekintjük sokaságnak, akkor a leírtak alapján látható, hogy ennek a sokaságnak jelentős része a 0 körül koncentrálódik. Emellett, például a vevőknél, megfigyelhető egy csomópont az 1 körül is, ugyanis a hibák jelentős része 100 százalék túlértékelés (például a már befolyt bevételt nem rendezték számvitelileg). A harmadik fontos probléma, hogy a legtöbb sokaság ferde, továbbá a ferdeség jellemző iránya és mértéke más és más az egyes beszámolóterületeken. A most felsorolt tulajdonságok miatt az általánosan használt eloszlások (normális, exponenciális, gamma, béta stb.) nem alkalmasak a valós és a könyv szerinti érték eltéréseinek modellezésére.

2. A valós érték pontbecslése

Tegyük fel, hogy n elemű egyszerű véletlen (a továbbiakban: EV-) mintát vettünk a sokaságból. Ez alapján egyebek mellett a következő módokon becsülhetjük a sokasági értékösszeget. Legegyszerűbb a mintaátlag alapján történő becslés:

$\hat{X}_m = N \cdot \frac{\sum x_i}{n}$. Ennek a becslésnek nyilvánvaló hátránya, hogy nem használja fel a pénzügyi kimutatásban szereplő, a valós értékkel jól korreláló adatokat.

A meglevő információt az $\hat{X}_d = Y - N\bar{d}$, az $\hat{X}_r = Y \cdot \left(\frac{\bar{x}}{\bar{y}}\right)$ különbség-, illetve hányadosbecsléssel, valamint az előző három valamilyen súlyozott átlagával használhatjuk fel.

Végül tételezzük fel, hogy az EV-minta helyett olyan visszatevés nélküli mintát vettünk, ahol minden sokasági elem mintába kerülési valószínűsége egyenesen arányos annak könyv szerinti értékével (a továbbiakban: PPS-minta, az angol probability proportional to size rövidítésből). Ilyen mintavételi terv mellett a sokasági értékösszegre adott torzítatlan Horvitz–Thompson-becslés

$$\hat{X}_{HT} = \sum \frac{x_i}{\left(\frac{n y_i}{Y}\right)}.$$

3. A valós érték intervallumbecslése

Az intervallumbecslés a mintavételi statisztikában szorosan összefügg a hipotézisvizsgálattal: azt az értéket nevezzük $100 \cdot \alpha$ százalékos ($\alpha \in [0;1]$) határnak,

amelyiket technikai nullhipotézisként vizsgálva a jobboldali próbánál a minta p -értéke éppen α . Amennyiben az intervallum két határa közül az egyik 0 vagy 100 százalék, egyoldali intervallumról beszélünk. Az $[\alpha_1; \alpha_2]$ határokkal definiált intervallumbecslés megbízhatósági szintje $\alpha_2 - \alpha_1$. Az ellenőrzési gyakorlatban alapvetően az egyoldali intervallumok terjedtek el, ezért a továbbiakban csak a $[0; 1 - \alpha]$ intervallummal, más néven a $100 \cdot (1 - \alpha)$ százalékos felső határral fogunk foglalkozni.

A legegyszerűbb módja az intervallumbecslésnek a pontbecslés mintavételi eloszlását veszi alapul, nevezetesen annak első két (centrális) momentumát, tehát az átlagot és a szórást. A tipikus becslési szituációkban tehát egy kétoldali intervallumbecslés a $\mu \pm \kappa_{1-\alpha/2} \cdot \sigma$ képlettel adható meg, ahol $\kappa_{1-\alpha/2}$ a pontbecslés standardizált eloszlásának megfelelő kvantilis értéke. Azokban az esetekben, amikor egy eloszlás „jól viselkedik”, a központi határeloszlás tétel alapján ezek a becslések már kisebb minták esetén is elfogadható eredményekre vezetnek.

A korábban leírt főbb statisztikai jellemzőkből kitűnik, hogy a könyvvizsgálat tipikus sokaságai nem követnek „jól viselkedő” eloszlást, és *Neter–Kim–Graham* [1975, 1977] vizsgálatai kimutatták, hogy ezeknél a sokaságoknál a hagyományos intervallumbecslési módszerek valóban jelentősen torzítanak. A torzítás egy része a ferdeségből, másik része a normális eloszlásétól eltérő lapultságból (csúcosságból) adódik, melyeknek az a folyománya, hogy a $\frac{\text{pontbecslés-valsó érték}}{\text{pontbecslés mintavételi szórása}}$ hányados még megközelítőleg sem követ t -eloszlást. (*Kaplan* [1973a, 1973b].)

Mivel a hagyományos módon készített becslések nem adtak kielégítő eredményt az eltérés nagyságára, a statisztika és a könyvvizsgálat határterületén több alternatív következtetési eljárást is kifejlesztettek, ezek egy része nagyban támaszkodik a sokasági aránybecslés módszereire.

A sokaság elemeiben található hiba eloszlását kevert eloszlással⁵ modellezzük: a hiba p valószínűséggel egy ξ valószínűségi változó értékeit ($E(\xi) = \theta, \xi \neq 0$) veszi fel, $1-p$ valószínűséggel pedig 0.⁶ $p_{1-\alpha}(m, n)$ jelöli M/N sokasági arány $1 - \alpha$ megbízhatóságú felső korlátját N elemű sokaság, n elemű minta, M minősített sokasági elem és m minősített mintabeli elem esetén (a sokasági arány becsléséről bővebben: *Lolbert* [2004]). A most bemutatandó becslések általános jellemzője, hogy az eltérés felső korlátját akarják megadni. Az alsó korláttal kapcsolatos lehetséges módosításokról a megfelelő helyen külön szólnunk.

⁵ Kevertnek nevezzük egy olyan valószínűségi változó eloszlását, amelynek értékeit úgy származtatjuk k darab különböző, előre rögzített eloszlásból, hogy a k -adik valószínűségi változó értékét pontosan p_k valószínűséggel veszi fel.

⁶ A kevert eloszlás meghatározását figyelembe véve a sokaságban található hiba eloszlását 2 valószínűségi változóból „kevertük ki”: egy tetszőleges olyan ξ valószínűségi változóból, mely $\xi(\omega) \neq 0 \quad \forall \omega$ esetén, és egy „determinisztikus” valószínűségi változóból, amely konstans 0 értékű.

3.1. Egy EV-mintán alapuló becslés

Ez a becslés feltételezi, hogy:

- a hiba túlértékelésből fakad ($Y_i \geq X_i$, azaz $D_i \geq 0$);
- a kimutatott tételek mind pozitívak ($Y_i > 0$);
- a tételben levő hiba maximális értéke legfeljebb a tétel értéke ($Y_i \geq D_i$, azaz $X_i \geq 0$).

Mindezeket figyelembe véve felírható a következő két reláció:

$$E(\xi) = \theta \leq Y_{\max},$$

(mivel ξ minden D_i realizációjára $D_i \leq Y_i \leq Y_{\max}$), illetve

$$E(D_i) = \sum_i \frac{D_i}{N} = p\theta \leq pY_{\max},$$

amiből egyszerű átalakítással a sokasági hiba összértékére kapjuk a $\sum D_i \leq NpY_{\max}$ felső korlátot. Amennyiben n elemű mintát veszünk a sokaságból, amelyben m hibás tételt találtunk, a

$$\hat{D}_{1-\alpha, EV} = Np_{1-\alpha}(m, n)Y_{\max} \quad /1/$$

becslésre fennáll a $Pr(\sum D_i \leq \hat{D}_{1-\alpha, EV}) \geq Pr(NpY_{\max} \leq \hat{D}_{1-\alpha, EV}) = 1 - \alpha$ reláció, tehát becslésünk *legalább* $100 \cdot (1 - \alpha)$ százalékban megbízható.

Figyeljük meg, hogy ez a becslés nem használja fel a mintában megfigyelt eltérések *nagyságát*, csupán a minta hibás tételeinek *arányát*, ezért *elvileg* jelentős pontosságjavulást lehet elérni egyrésztől rétegzett mintavétellel, másrésztől a maximális hibanagyságra tett feltevés módosításával. A gyakorlatban ennek ellenére ezt a módszert ritkán használják, alapvetően azért, mert szinte kivétel nélkül PPS-elvű (ezen belül is MUS – monetary unit sampling – pénzegység alapú) mintavételt alkalmaznak.⁷

3.2. A PPS-mintán alapuló becslések

A beszámolóok auditálásakor az egész világon széles körben használt, gyakorlatban előforduló becslési eljárások legfontosabb közös jellemzője a MUS-, vagy DUS-

⁷ Vegyük azonban észre, hogy a PPS-mintavétel a rétegzett mintavétel speciális határeseteként értelmezhető.

mintavétel (monetary unit sampling vagy dollar unit sampling, a hazánkban elterjedt terminológia szerint pénzegység alapú mintavétel). A MUS a könyvvizsgálói gyakorlatban olyannyira elterjedt és elfogadott módszer lett, hogy szinte mást nem is használnak, és általában figyelmen kívül hagyják a módszer meglévő korlátjait, előfeltételeit, így sokszor azokra a következtetésekre is alkalmazzák, amikre alkalmatlan.

A MUS valójában csak annyit jelent, hogy a mintát az eredeti sokaság pénzegységeiből alkotott mesterséges sokaságból veszik, majd megvizsgálják azokat az eredeti tételeket, amelyekből pénzegységet választottak.⁸ Könnyen bizonyítható, hogy ez a kiválasztási módszer az eredeti sokaságra nézve egy PPS-mintát eredményez. MUS-mintavétel esetén a hipotetikus sokaság „tételszáma” (elemszáma) Y (az *eredeti sokasági értékösszeg*), a tételek (sokasági elemek) „könyv szerinti értéke” pedig definíció szerint a hipotetikus sokaság minden egyes elemére 1. A minta elemszámához (n) képest Y gyakorlatilag végtelennek tekinthető (pár száztól több-millió/milliárdig terjedhet), ezért mindegy, hogy visszatevéssel vagy visszatevés nélkül veszünk-e mintát.

Ennek a mintaválasztási megközelítésnek a könyvvizsgálatban való alkalmazására tett első utalás még 1961-ből, *van Heerden* holland nyelvű cikkéből (*van Herden* [1961]) származik, de a könyvvizsgálói szakma szélesebb köre csak 1963-ban ismerte meg, *Kenneth W. Stringertől* (*Stringer* [1963]). Az általa akkor még csak nagy vonalakban leírt MUS-módszer egyik legismertebb becslési eljárását *Stringer-féle felső határnak* (*Stringer bound*) hívják. A hetvenes években több alternatív módszert is kifejlesztettek, melyek legtöbbször azonban továbbra is magán hordozza a később bemutatandó *Stringer-féle* becslés gyengeségeit: a becslések torzítatlansága analitikusan nem igazolható, a szimulációk alapján pedig a becslések jó része túlságosan konzervatív, azaz a névleges szintnél jóval magasabb a megbízhatóságuk, és így jóval kisebb a pontosságuk (túlságosan széles az intervallum). A MUS-mintát használó módszerek általában (így az itt leírásra kerülő *Stringer*-, *cella*- és *multinomiális* módszerek is) az úgynevezett *CAV*- (*combined attributes-variables sampling*) elven alapulnak. A *CAV*-becslések „diszkrétte teszik” az eredeti eloszlást olyan módon, hogy az eltéréseket nagyságuk alapján intervallumokba (kategóriákba) sorolják, és helyettesítik őket az intervallum egyik (jellemzően a legnagyobb) értékével.

A MUS-minta kiválasztásának technikai lebonyolítása

A pénzegység alapú mintavételt technikailag többféleképpen lehet elvégezni, melyből a következő módszereket érdemes kiemelni.

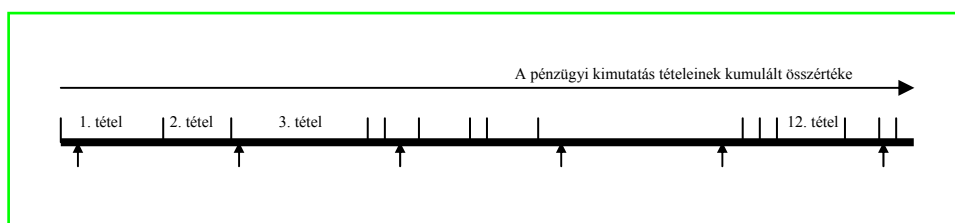
1. Az elmélet szempontjából legegyszerűbb esetben a pénzegységekből ún. korlátozás nélküli mintát, azaz *EV*-mintát veszünk. Ebben

⁸ A MUS a statisztikában ismert „kumulált értékösszegek módszere” egy alkalmazásának is tekinthető.

az esetben akár az is előfordulhat, hogy minden n alkalommal ugyanazt a tételt kell megvizsgálunk, és emiatt a korlátozás nélküli MUS-mintavétel sokak számára nem elfogadható. Ennek ellenére ezt a mintavételi technikát tekintjük alapértelmezésnek a továbbiakban, számtalan jó tulajdonsága miatt.

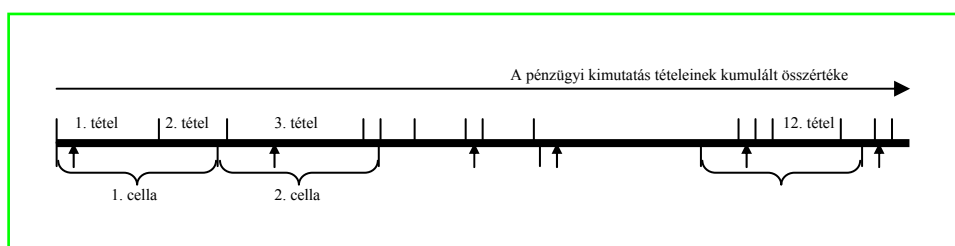
2. A tételek valamilyen előre rögzített sorrendben kumulált sorozatának egy véletlenszerűen kiválasztott pontjáról elindulva n alkalommal felmérünk Y/n nagyságú lépésközt. Ez a manuális gyakorlatban leginkább elterjedt módszer, a szisztematikus kiválasztási módszerekre jellemző egyszerűségének köszönhetően. A könnyebb megértés kedvéért tekintsük az 1. ábrát.

1. ábra. Tételek a MUS-mintában



Az 1. ábrán a felső beosztás intervallumai jelzik az egyes tételeket. Az első nyíl mutatja a véletlenszerűen kiválasztott pontot, a két szomszédos nyíl közötti távolság pedig a lépésközt. Vegyük észre, hogy az 1. ábrán a 4. és az 5. nyíl ugyanazt a tételt jelöli meg. Az ilyen tételeket nevezik lépésköz feletti, vagy nagy értékű tételeknek (HVI – high value items). A MUS egyes változatai a nagy értékű tételeket más és más módon kezelik, de jelen tanulmány szempontjából ez nem lényegi kérdés.

2. ábra. Cellák a MUS-mintában



3. Az úgynevezett cella-módszerben a 2. pontban leírttal szemben a tételek kumulált sorozatát n darab, Y/n hosszúságú intervallumra („cel-

lára”) osztjuk, és minden egyes intervallumon belül véletlenszerűen kiválasztott 1-1 pont (az ábrán továbbra is nyíllal jelölve) határozza meg a mintaelemeket. A módszerhez külön kiértékelő formula is tartozik, aminek részleteiről külön alpontban fogunk írni.

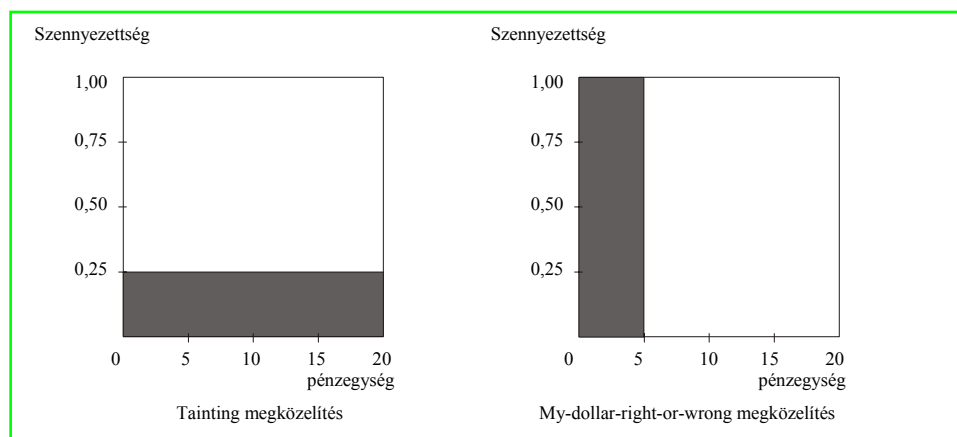
A három mintavételi terv a tételek rögzített sorrendje esetén nem egyenértékű. Vegyük észre, hogy sem a 2., sem a 3. terv nem képes produkálni minden olyan mintát, amit az 1. módszer eredményezhet: sem a 2., sem a 3. mintavételi terv nem tud például olyan mintát eredményezni, amiben mind a 6. mind a 7. tétel szerepel. Hasonló módon a 2. terv sem képes minden olyan mintát produkálni, amit a 3. tud. Könnyen látható azonban, hogy a tételek mintavétel előtti „megkeverésével” (véletlenszerű permutálásával) ez a különbség megszűnik. Nehéz analitikusan átlátni, hogy a mintavételi terveknek ez a különbsége pontosan milyen hatást gyakorol egy adott kiértékelő formulára, és ezzel kapcsolatosan az általam ismert irodalom sem nyújtott kellő mértékű eligazítást.

További problémát okoz ezeknél a mintavételi terveknél annak eldöntése, hogy a mintába választott különböző pénzegységek milyen mértékű hibát tartalmaznak. Ezzel kapcsolatosan két felfogás létezik. Az uralkodó, de időben későbbi megközelítés (tainting-elv) szerint a pénzegység hibája az őt tartalmazó fizikai tétel szennyezettségével egyezik meg, tehát bárhonnét vesszük ki az adott tételből a mintaelemet, a hiba ugyanaz. A másik megközelítés (my-dollar-right-or-wrong) az adott fizikai tétel hibáját a tétel elejétől (egyedülálló alkalmazásokban a végétől) kezdi számolni, tehát attól függően, hogy honnét származik a mintaelem, a hibája 1, 0, vagy pedig egy tört (pont a határon van, és az eredeti tételben szereplő hiba nem egész szám forintban nézve). Ha tehát a példa kedvéért egy tétel 25 százalékos szennyezettségű, és 20 pénzegységből áll, akkor a tainting-elv szerint minden egyes pénzegység 25 százalékos szennyezettségű. Ezzel szemben a másik megközelítésben az első 25 százalékos (az első 5 egységet) 100 százalékosan szennyezettnek, a továbbiakat viszont teljesen szennyezettségmentesnek tekintik.⁹ Ezt a helyzetet a 3. ábra szemlélteti.

Érdekes módon a tainting megközelítés – annak ellenére, hogy kismintás tulajdonságai jobbak, aszimptotikusan alulmarad a my-dollar-right-or-wrong megközelítéssel szemben (lásd például *Pap-van Zuiljen* [2000]). A továbbiakban mi a tainting megközelítést fogjuk alkalmazni. A mintaválasztási módok és a hibák különböző értékelései közötti eltérések jobb megértésére egy leegyszerűsített példán kövessük végig alkalmazásukat.

⁹ A my-dollar-right-or-wrong megközelítés gyakorlatilag nem más, mint a sokasági aránybecslés közvetlen alkalmazása a pénzegységek mesterséges populációjára.

3. ábra. A pénzegységek szennyezettsége



Először tekintjük a következő elszámolást,¹⁰ amit egy külföldi kiküldetésből hazatért kolléga nyújtott be.

1. táblázat

Külföldi kiküldetés elszámolása

Sorszám	Megnevezés	Könyv szerinti érték (Y_i)	Valós érték (X_i) (a priori ismeretlen)	Megjegyzés	Relatív hiba (százalék)
		euró			
1	Taxiszámla a Ferihegyi repülőtérre	23	0	Kapott BKV bérletet, így nem jogosult elszámolni	100
2	Repülőjegy oda-vissza	512	512		0
3	Taxi a szállásig	72	72		0
4	Szállás 5 napra félpanzióval.	432	324	Elírás	25
5	Helyi tömegközlekedés, heti bérlet	84	84		0
6	Éttermi ebéd 1. nap	15	15		0
7	Éttermi ebéd 2. nap	15	15		0
8	Éttermi ebéd 3. nap	15	15		0
9	Éttermi ebéd 4. nap	15	15		0
10	Éttermi ebéd 5. nap	15	15		0
11	Telefonköltség	43	43		0
12	Taxi a repülőtérre	68	68		0
13	Taxi a Ferihegyi repülőtérről	35	0	Lásd. 1. tételnél	100
<i>Összesen</i>		1344	1178		12

¹⁰ Mindössze 13 tétel esetén a valós alkalmazásokban természetesen nincs mintavétel, hanem teljes körű tételes ellenőrzést alkalmaznak.

A tételek közül szűrőpróbaszerűen kiválaszt a pénzügyes 6 tételt. Ehhez először el kell készíteni a tételek kumulált sorozatát.

2. táblázat

Tételek kumulált sorozata

Sorszám	Könyv szerinti érték	Kumulált könyv szerinti érték
	(euró)	
1	23	23
2	512	535
3	72	607
4	432	1039
5	84	1123
6	15	1138
7	15	1153
8	15	1168
9	15	1183
10	15	1198
11	43	1241
12	68	1309
13	35	1344

1. Az első mintavételi módszer alkalmazásához 6 elemű EV-mintát veszünk az 1, 2, ... 1344 számokból. Legyenek ezek a 17, 52, 364, 836, 1293 és 1317. Ezt felhasználva a mintánk az 1., 2., 2., 8., 12. és 13. tételekből áll. Látható, hogy a 2. tétel kétszer szerepel a mintában.

2. A második mintavételi módszerhez két értéket kell meghatározni: a lépésközt és a kezdőpontot. A lépésköz Y/n , azaz $1344/6=224$, a kezdőpont pedig az 1, 2, ... 1344 számokból választott 1 elemű EV-minta, ami ez esetben legyen mondjuk 3. A mintába eső pénzegységek a 3, $3+224=227$, $3+224+224=451$, 675, 899, 1123. Fizikai tételekre lefordítva ez az 1., 2., 2., 4., 4., 5. tétel. A 2. és a 4. tétel HVI, ezért mindenképpen a mintába kellett kerülniük. A 2. tétel nagysága a lépésköz kétszeresét is meghaladja, ezért mindenképpen kétszer kerül a mintába. A 4. tétel bizonyos kezdőpontok esetén egyszer, bizonyos kezdőpontok esetén kétszer kerül a mintába.

3. A cellamódszerhez először meg kell határozni a cellákat. A cellák nagysága megegyezik a lépésközzel, így a cellahatárok 1–224,

225–448, 449–672, 673–896, 897–1120, 1121–1344. Ezek után 6 elemű visszatevéses mintát veszünk az 1–224 intervallumból, legyen ez 27, 143, 53, 197, 81, 152. A kiválasztott pénzegységek az $1+27-1=27$, $225+143-1=367$, $449+53-1=501$, $673+197-1=869$, $897+81-1=977$, $1121+152-1=1272$. Az ezekhez tartozó tételek sorszáma rendre 2, 2, 2, 4, 4, 12. A cellamódszerben a szerencsétlen véletlenek miatt egy tétel annyiszor kerülhet a mintába, ahány cellával van metszéspontja. Ezeket az ismétléseket a cellamódszer egy fejlettebb megvalósítása kiszűri, de ennek ismertetése túllépi a tanulmány kereteit.

A két hibamérési megközelítés közötti különbséget a második módszerrel választott mintán mutatjuk meg, konkrétan a 4. elszámolt tétel esetén. Emlékezzünk, hogy ez a tétel kétszer került a mintába. A 4. tétel a 608. pénzegységtől az 1039. pénzegységig tartott.

1. Az első felfogás, a tainting-elv szerint a mintába került 432 eurból 75 százaléknyi helyes, 25 százaléknyi hibás, mindkét esetben $(324/432=0,75=75\text{ százalék})$.

2. A hibát a tétel elejétől felmérő my-dollar-right-or-wrong-felfogás szerint a hibás pénzegységek a 608-tól 715-ig tartanak $((1039-608)*0,25+608=715)$. Így a mintába került első pénzegység (675) hibája 100 százalék (mert $675 < 715$), míg a második pénzegység (899) hibája 0 százalék (hiszen $899 > 715$).

3. A hibát a tétel végétől felmérő my-dollar-right-or-wrong-felfogás szerint a 931-től 1039-ig található pénzegységek a hibásak. Így a mintába került mindkét pénzegység 0 százalék hibát tartalmaz.

Az EV-becslés alkalmazása a mesterséges sokaságra

Az EV-mintán alapuló /1/ becslést alkalmazva a MUS mesterséges sokaságára a következőt kapjuk (Y darab tétel, a tételek „könyv szerinti értéke” pedig definíció szerint 1):

$$Pr\left(\sum D_i \leq Yp_{1-\alpha}(m, n)\right) \geq 1 - \alpha.$$

Mivel az eredeti becslésnél $Y \leq NY_{max}$, ezért a MUS-mintán alapuló

$$\hat{D}_{1-\alpha, MUS} = Yp_{1-\alpha}(m, n) \quad /2/$$

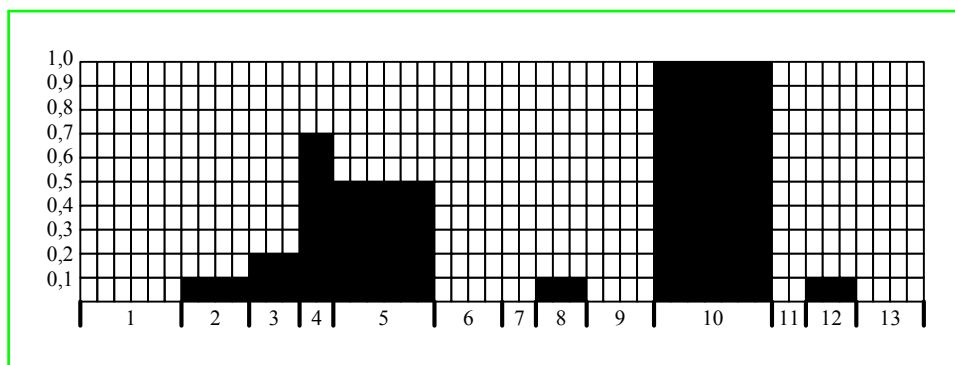
becslés jóval pontosabb (kevésbé konzervatív, szűkebb az intervallum) az EV-mintán alapuló becslésnél. Mindazonáltal ez a becslés sem veszi figyelembe, hogy nem minden hibás tétel 100 százalékgig hibás, így ez a becslés is túlságosan óvatosnak tekinthető.

Az eddig leírt két elemi módszernek minden hibájuk ellenére megvolt az a jó tulajdonsága, hogy a megbízhatóságuk analitikusan igazolható. A most bemutatandó becslésekre ez már sajnos nem, vagy csak korlátozottan lesz igaz. Ezeknél a kapcsolódó irodalom szinte kivétel nélkül szimulációkkal igyekszik a megbízhatóságról, illetve a torzítás mértékéről meggyőződni.

A Stringer-féle felső határ (Stringer bound)

Mielőtt leírnánk a Stringer-becslést részleteiben, tekintsük meg a 4. ábrát.

4. ábra. A könyv szerinti érték szennyezettsége



A 4. ábrán látható kis négyzet magassága 0,1 Ft, szélessége 1 Ft, tehát a rács egy oszlopa jelenti a sokaság egy elemét. Az eredeti sokaságot az alsó, vastagabb vonalak definiálják, tehát az eredeti sokaság könyv szerinti értékei rendre 6, 4, 3, 2, 6... stb. Mivel ez esetben feltételezzük, hogy a valós érték nem negatív, de legfeljebb a könyv szerinti érték, ezért egy adott tétel valós értékét úgy ábrázoljuk, hogy befejejtjük a könyv szerinti érték szennyezettség mértékének megfelelő hányadát. Ez alapján például a második tétel könyv szerinti értéke 4, valós értéke 3,6; a 10. tétel pedig nem ér semmit a valóságban.

A jelenséget vizuálisan megközelítve, a pénzügyi kimutatásban található összes hibák becslése ugyanaz a probléma, mintha a Szahara egy szabályos téglalap alakú részén található felszíni vizek összterfogatára akarnánk úgy felső becslést adni, hogy véletlenszerűen kiválasztott GPS-koordinátáknál ismerjük a vízmélységet (ez a legtöbb helyen tipikusan 0 lesz).

A korábban leírt /2/ becslést felírhatjuk a következő alakban is:

$$\hat{D}_{1-\alpha, MUS} = Y p_{1-\alpha}(m, n) = 10 \cdot 0,1 \cdot Y p_{1-\alpha}(m, n) = \sum_{i=1}^{10} 0,1 \cdot Y p_{1-\alpha}(m, n).$$

A 4. ábrában ez azt jelenti, hogy külön számoltunk minden 10 fillérre felső korlátot, amiket aztán összeadtunk. Ha pontosítani szeretnénk a becslést, az első intuíció ebből a képletből kiindulva azt sugallhatja, hogy nézzük meg, hány mintába került tételben van *legalább* 1, 2, ... 100 fillér hiba (ez nyilván egy monoton csökkenő sorozat lesz, ugyanis ha $a > b$, akkor a legalább b hibát tartalmazó tételek halmaza tartalmazni fogja a legalább a hibát tartalmazó tételek halmazát. A két halmaz különbsége a *pontosan* $b+1, b+2, b+3 \dots a$ mennyiségű hibát tartalmazó tételek uniója). Ezután külön-külön adjunk felső becslést az adott kategória (tehát például legalább 2 fillér hiba stb.) sokasági arányára, és ezeket a kategóriák nagyságával (ebben az esetben $Y/100$) súlyozva adjuk össze. Amint nemsokára látni fogjuk, ez egy, a Stringer-féle felső határhoz nagyon hasonló felső határt fog megadni. A gondolatmenetben azonban egy súlyos hiba van: nem igaz ugyanis, hogy $Pr(\xi \leq x) = 1 - \alpha$ és $Pr(\eta \leq y) = 1 - \alpha$ relációkból következne $Pr(\xi + \eta \leq x + y) = 1 - \alpha$ reláció, csupán $Pr(\xi + \eta < x + y) \in [1 - 2\alpha; 1]$ állítható. A Stringer-féle felső határ egy rendezett mintás statisztika: a nagyság szerint csökkenő szennyezettségeket rögzített (csak a mintamérettől függő), a helyezésüknek megfelelő súlyokkal „átlagoljuk”, majd ezt az „átlagot” megszorozzuk a teljes sokaság könyv szerinti értékével. Képlettel:

$$\hat{D}_{1-\alpha, st} = Y \cdot \left(p_{1-\alpha}(0, n) \cdot 1 + \sum_{i=1}^n [p_{1-\alpha}(i, n) - p_{1-\alpha}(i-1, n)] \cdot t_i \right), \quad /3/$$

ahol t_i a mintában található i -edik legnagyobb szennyezettséget jelenti. Ezt képletet átalakítva kapjuk:

$$\begin{aligned} \hat{D}_{1-\alpha, st} &= Y \cdot \left(p_{1-\alpha}(0, n) \cdot 1 + \sum_{i=1}^n [p_{1-\alpha}(i, n) - p_{1-\alpha}(i-1, n)] \cdot t_i \right) = \\ &= Y \cdot \sum_{i=0}^n (t_i - t_{i+1}) \cdot p_{1-\alpha}(i, n), \end{aligned} \quad /4/$$

ahol $t_0 = 1$ illetve $t_{n+1} = 0$.

A 4. ábrán 100, 70, 50, 20, 10 és 0 százalékos szennyezettségeket látunk, tehát ez a képlet rendre felső becslést ad a legalább 100, 70, 50, 20, 10 százalékos hibák arányára, ezeket 100-70, 70-50, 50-20, 20-10, 10-0 százalékos súlyokkal súlyozza, majd

ezt a súlyozott összeget (átlagos hibamértéket) kiveti a teljes sokaságra, Y -ra. Ha ebben a formában nézzük tehát a Stringer-féle felső határt, azonnal látszik, hogy ebben az esetben is az intervallumokba esés valószínűségének felső határát becsültük meg, és elköveltük azt a már említett hibát, hogy ezeket az értékeket mechanikusan összeadtuk. Természetesen ez még nem jelenti azt, hogy a Stringer-bebecslés nem lenne jó, de ezzel a megközelítéssel jósága nem bizonyítható.

Amikor az eredeti képlet megjelent, még nem állt olyan számítástechnikai háttér rendelkezésre, mellyel gyorsan meghatározható lett volna $p_{1-\alpha}(m, n)$ értéke tetszőleges N , m , n és megbízhatósági szint esetén. Mivel nagyobb sokaságokra a hipergeometriai eloszlás közelíthető binomiálissal, illetve alacsony hibaarányok mellett a binomiális eloszlás közelíthető a könnyen kezelhető Poisson-eloszlással, ezért kezdetben a $p_{1-\alpha}(m, n)$ érték helyett a $\frac{\lambda_{1-\alpha}(m)}{n}$ közelítő értéket használták, ahol $\lambda_{1-\alpha}(m)$ a Poisson-eloszlás paraméterére vonatkozó $100 \cdot (1 - \alpha)$ százalékos felső határ m megfigyelt hiba mellett. Ennek a közelítésnek megvolt az az előnye, hogy a rendezett mintás statisztika súlyait táblázatba lehetett gyűjteni a mintamérettől függetlenül, hiszen az kiemelhető volt a képletből:

$$\hat{D}_{st} = \frac{Y}{n} \cdot \left(P_0 + \sum_{i=1}^n P_i \cdot t_i \right),$$

ahol Y/n a lépésköz (itt vesszük figyelembe a mintaméretet), P_i pedig a táblázatból kiolvasható i -edik úgynevezett Poisson-faktor, tehát $P_i = \lambda(i) - \lambda(i-1)$. (Az irodalom nem teljesen következetes, egyes szerzők $\lambda(i)$ -t nevezik Poisson-faktornak.) Valószínűleg egyébként éppen ezért a könnyű kezelhetőségért terjedt el ebben a formájában a képlet, és nem a másik, $p_{1-\alpha}(m, n)$ szerint csoportosított formában. Noha ma már bármely korszerű személyi számítógép azonnal ki tudná számolni a hipergeometriai faktorokat is, a Poisson-közelítéssel való számolás még mindig nagyon elterjedt a könyvvizsgálók között.

Mivel a Stringer-sejtés (tehát hogy a becslőfüggvény *legalább* $100 \cdot (1 - \alpha)$ százalékban megbízható) általános feltételek mellett mindeztidáig nem került igazolásra, és ellenpéldát sem sikerült konstruálni, számos szimulációt végeztek és publikáltak a témában. A szimulációk erős empirikus bizonyítékot szolgáltatottak arra, hogy a becslés megbízhatósága jóval meghaladja a névleges $100 \cdot (1 - \alpha)$ százalékot (például az általánosan használt 95 százaléknál az esetek 98-99 százalékában haladta meg a becsült felső határ a valós értéket).

Függetlenül azonban attól, hogy mekkora a becslés megbízhatósága, fennáll a $\hat{D}_{1-\alpha, st} \leq \hat{D}_{1-\alpha, MUS}$ reláció minden olyan esetben, amikor a szennyezettségek 0 és 1 közé esnek, ugyanis:

$$\hat{D}_{1-\alpha, st} = Y \cdot \sum_{i=0}^n (t_i - t_{i+1}) \cdot p_{1-\alpha}(i, n) \leq Y \cdot \sum_{i=0}^n (t_i - t_{i+1}) \cdot p_{1-\alpha}(m, n) = \hat{D}_{1-\alpha, MUS},$$

mivel $t_0 = 1$ és m az a legkisebb egész szám, amelyre $t_m \leq 0$ (a szennyezettségek monoton csökkennek és nem negatívak). Egyenlőség áll fenn, ha a szennyezettségek csak 0 vagy 1 értéket vehetnek fel. Ennek alapján kijelenthető, hogy amennyiben *legalább* $100 \cdot (1 - \alpha)$ százalékban megbízható az EV-mintán alapuló becslés, akkor a pontossága jobb, mint a MUS-mintán alapuló becslésnek.

Ha a Stringer-féle felső határ /4/ alatti alakjából (illetve az alakhoz kapcsolódó intuíciónál) közelítünk, akkor a szumma első tagja veszi számba a mintában ugyan nem található, de feltételezhetően meglévő szennyezettséget. Az eredeti képletben ez $t_0 = 1$, azonban számos területen a gyakorlati tapasztalatok szerint jóval kisebb az elképzelhető legnagyobb hiba. Amennyiben tehát biztos információval rendelkezünk az elképzelhető legnagyobb hiba nagyságáról (például korábbi ellenőrzések alapján, vagy az intézményrendszer ismeretében), akkor ezt az értéket t_0 helyébe állítva jelentősen élesíthetjük a becslésünkön (az így kapott becslés neve: generalized Stringer bound, tehát általánosított Stringer-féle felső határ). A legfontosabb könyvvizsgálati szoftverekben is állítható ez az érték, általában BPP (basic precision pricing) a neve. Az elnevezés onnét származik, hogy nagyságrendileg általában ez a paraméter befolyásolja leginkább a becslésünket, és nem a mintából származó szennyezettségek.

Noha a Stringer-sejtést teljes egészében eddig nem igazolták, több fontos rész-eredmény született. Az első, úttörőnek tekinthető írás *Bickel* [1992] tanulmánya. A szerző bizonyítja, hogy ha:

– ξ folytonos valószínűségi változó, akkor

$$P(D \leq \hat{D}_{1-\alpha, st}) \geq (1 - \alpha)^{n+1}, \text{ illetve ha}$$

– ξ legfeljebb 2 értéket vehet fel, akkor

$$P(D \leq \hat{D}_{1-\alpha, st}) \geq 1 - \alpha,$$

tehát a becslés *legalább* $100 \cdot (1 - \alpha)$ százalékban megbízható.

A cikk további fontos eredménye, hogy a becslés felső korlátot fel tudja írni ξ várható értékének, illetve az eloszlásfüggvény egy bonyolult integráljának összegeként. Ennek a felírásnak a jelentősége, hogy segítségével kiszámítható a becslés aszimptotikus (végtelen mintaméretnél értelmezett) eloszlása.

Pap, van Zuijlen és de Jager több tanulmányban [1995, 1996, 1997] folytatja a Bickel által elkezdett megközelítést. Három legfontosabb eredményük a következő.

1. A Bickel által felírt aszimptotikus eloszlás segítségével bizonyítják, hogy $\alpha \in [0; 0,5]$ esetben a becslés megbízható, ellenkező esetben aszimptotikusan nem megbízható. A valós megbízhatósági szint az első esetben jóval meghaladja, a második esetben viszont még közelítőleg sem éri el a névleges $100 \cdot (1 - \alpha)$ százalékot. Mivel az auditorok általában 50 százalék feletti megbízhatósági szinttel dolgoznak, ezért csak az $\alpha \in [0; 0,5]$ esetre tett megállapításoknak van gyakorlati jelentősége.

2. Az előbbi észrevételt kiegészítve bevezetnek egy olyan módosított becslőfüggvényt, amely aszimptotikusan pontosan a névleges szintnek megfelelő megbízhatóságú.

3. A szennyezettségek tetszőleges olyan eloszlása esetén, ahol a szennyezettségek csak 0 és 1 értéket vehetnek fel, minden olyan lehetséges $a_{1-\alpha}(i, n) \geq a_{1-\alpha}(i-1, n)$ együttható-sorozatra, amelyre a

$$\hat{D}_{1-\alpha} = Y \cdot \left(a_{1-\alpha}(0, n) \cdot 1 + \sum_{i=1}^n [a_{1-\alpha}(i, n) - a_{1-\alpha}(i-1, n)] \cdot t_i \right)$$

felső határ legalább $100 \cdot (1 - \alpha)$ százalékban megbízható, $a_{1-\alpha}(i, n) \geq p_{1-\alpha}(i, n) \forall i$ -re. Ez azt jelenti, hogy ebben az esetben bizonyos tekintetben minimálisak a képletben szereplő együtthatók.

A Stringer-féle felső határral kapcsolatos irodalomban gyakorlati szempontból óriási jelentőségű *Neter–Kim–Graham* [1984] tanulmánya. A gyakorlatban ugyanis sokszor olyan nagy az auditálandó beszámoló, hogy azt részterületre bontva lehet csak vizsgálni. (Ilyen volt például a tanulmány elején az Eszközök összesen értékének vizsgálata.) Az egyes részterületeken egymástól függetlenül történik a mintavétel, ennek ellenére véleményt kell mondani a teljes beszámoló megbízhatóságáról is. A legnagyobb problémát az jelenti, hogy a részterületek különbözősége miatt a teljes beszámolóra nézve már nem áll fenn, hogy a mintába kerülés valószínűsége minden tételre arányos lenne a tétel beszámolóban szereplő nagyságával. Az említett tanulmány feltételezi, hogy az egyes részpopulációkból független MUS-mintát vettek, amiket a Stringer-féle felső határt használva értékelték ki. A szerzők a kombinált felső határ kiszámolására 5 különböző megoldást javasolnak, melyek a következők.

1. *Független valószínűségi változók összeadása.* Feltételezve, hogy az egyes részterületekre számított felső határok függetlenek egymástól,

a teljes beszámoló felső határának megbízhatósági szintje legalább $\prod_i (1 - \alpha_i)$, ahol $1 - \alpha_i$ az i -edik részsokaságra tett felsőhatár-becslés megbízhatósága. Ez alapján ha minden részsokaságra egyforma megbízhatósági szintet használunk, az aggregált becsléshez elvárt legalább $1 - \alpha$ megbízhatóság a részsokaságoknál legalább $(1 - \alpha)^{1/L}$ megbízhatóságú becsléseket igényel, ahol L a részsokaságok száma. (Tehát 95 százalékos megbízhatósághoz 2 részterület esetén mindkét részterületen 97,5 százalékos megbízhatóságú becslést kell készíteni.)

2. *Implicit standard hiba használata.* Mivel egy adott mintánál nem csak az adott megbízhatóságú egyoldali intervallum végpontja meghatározható, hanem a pontbecslés is, ezért implicit módon, „visszafelé” kiszámíthatjuk azt a standard hibát, amit a normális eloszlással való közelítés esetén használva ugyanezt a felső határt kaptuk volna. Formálisan:

$$SE = \frac{\hat{D}_{1-\alpha} - \hat{D}}{z_{1-\alpha}}.$$

Ezt a standard hibát használva kiszámolhatjuk a független változók összegének standard hibáját is, amiből a megszokott módon (pontbecslés + $z \cdot SE$) kapjuk az összegre vonatkozó becsült felső határt.

3. *Közelítő globális kiértékelés.* A felsőhatár-becslés

$$\hat{D}_{1-\alpha, st} = Y_l \cdot \left(p_{1-\alpha}(0, n_l) \cdot 1 + \sum_{i=1}^n [p_{1-\alpha}(i, n_l) - p_{1-\alpha}(i-1, n_l)] \cdot t_i \right)$$

képletében részsokaságonként különböző sokasági értékek (Y) és mintanagyságok (n) szerepelhetnek. Közelítő globális kiértékelés esetén a zárójelben szereplő első tag (az ún. basic precision) nélkül vesznek figyelembe minden sokaságot, kivéve azt a sokaságot, amelynél a legnagyobb az $Y_l \cdot p_{1-\alpha}(0, n_l)$ szorzat értéke. Képlettel:

$$\hat{D}_{1-\alpha, st}^{kombinált} = \max_l (Y_l \cdot p_{1-\alpha}(0, n_l)) + \sum_l (\hat{D}_{1-\alpha, st, l} - Y_l \cdot p_{1-\alpha}(0, n_l)).$$

4. *Globális kiértékelés.* A globális kiértékelés egy nagy mintának kezeli a részsokaságokból származó L különböző mintát, és erre a mintára alkalmazza a Stringer-becslés egy módosított változatát (az „összefésült” mintaelemeket itt is nagyság szerint csökkenő sorrendbe rakjuk):

$$\hat{D}_{1-\alpha, st}^{kombinált} = \max_l (Y_l \cdot p_{1-\alpha}(0, n_l)) + \sum_i [p_{1-\alpha}(i, n_l) - p_{1-\alpha}(i-1, n_l)] \cdot t_i \cdot Y_l,$$

ahol az l alsó index annak a mintának a nagyságára és kimutatott értékére utal, amiből az adott szennyezettség származik.

5. „Konzervatív” globális kiértékelés. A konzervatív változat annak a részmintának az n és Y paramétereit használja mindenhol, amelyre $Y_l \cdot p_{1-\alpha}(0, n_l)$ felveszi a maximumát. Azzal a kérdéssel a tanulmány nem foglalkozik, mi történik több ilyen részminta esetén, ugyanis akkoriban még az elméleti munkákban is a Poisson-közelítést használták, és a közelítő felírásban ennek a kérdésnek nincs jelentősége.

A tanulmányban szereplő 5 kiértékelési módszerre vonatkozóan több szimulációt is végeztek a szerzők, melyek kivétel nélkül 98-99 százalékos megbízhatóságot mutattak a kombinált felső határra 95 százalék elvárt megbízhatóság mellett.

A cellamódszer (Cell bound)

A cellamódszernél leírt mintaválasztási módszerhez a szerzők (*Leslie, Teitlebaum, Anderson* [1980]) külön kiértékelési metódust dolgoztak ki. Mivel a módszer jóval kevésbé intuitív, mint akár a Stringer-féle felső határ, akár következő szakaszban ismertetésre kerülő multinomiális felső határ, ezért most csupán a kiértékelés módját írjuk le, intuitív indoklás nélkül. A most következő leírás megegyezik az eredetileg leírtakkal, így a Poisson-eloszlással közelíti a valószínűségeket. Az elmúlt években a legtöbb gyakorlati alkalmazás (többek között az IDEA szoftver is) már az egzakt hipergeometriai faktorokat használja.

A mintán megfigyelt szennyezettségeket ebben az esetben is csökkenő sorrendbe állítjuk, és ezután alkalmazzuk a következő rekurzív formulát:

$$F(0) = \lambda_{1-\alpha}(0),$$

$$F(i) = \max(F(i-1) + t_i, \lambda_{1-\alpha}(i) \cdot t_i), \text{ egészen az utolsó hibáig, } m\text{-ig.}$$

A becsült felső határ a legutolsó F és lépésköz, Y/n szorzata:

$$\hat{D}_{1-\alpha, cell} = F(m) \cdot \frac{Y}{n}.$$

A cellamódszer kiértékelő része a szimulációk alapján kevésbé konzervatív, mint a Stringer-féle felső határ, azonban megbízhatósága még így is jóval meghaladja a

névlegest. Előnye, hogy a Poisson-faktorokat tartalmazó táblázat segítségével számítógép nélkül is meghatározható.

Multinomiális felső határ (Multinomial bound)

A multinomiális módszert *Fienberg, Neter és Leitch* 1977-ben publikálta (*Fienberg–Neter–Leitch* [1977]), tehát gyakorlatilag egy időben került kifejlesztésre a *Leslie, Teitlebaum, Anderson*-féle cellamódszerrel (*Leslie–Teitlebaum–Anderson* [1980]). A cellamódszerrel szemben ez a becslés csak számítógép segítségével alkalmazható.

A multinomiális modell eredetileg leírt változatában minden pénzegységet besorolnak 101 kategória valamelyikébe aszerint, hogy az adott pénzegységre jutó szennyezettség mértéke 0 százalék, 0 százaléknál több de legfeljebb 1 százalék, 1 százaléknál több de legfeljebb 2 százalék és így tovább 100 százalékig. Ha a sokaságban az i -edik csoportba eső elemek aránya p_i , akkor a sokasági hibarányra $\sum_{i=0}^{100} \frac{i}{100} \cdot p_i$

felső becslést ad, és a becslés legfeljebb 1 százalékponttal haladja meg a valós értéket. Amennyiben a mintába kerülő pénzegységeket visszatevéssel választjuk, vagy pedig a minta mérete a sokaságéhoz képest elhanyagolható, akkor a mintaelemek 101 kategória közötti eloszlása (n, p_i) paraméterű multinomiális eloszlást követ, ahol az első paraméter a mintaméret, a további 101 darab p_i paraméter pedig az indexe által meghatározott csoportba való esés valószínűsége. A multinomiális felső határ megadásához két lépés vezet.

Az első lépésben bevezetünk egy rendezést a lehetséges minták között, melynek segítségével meghatározhatók a kapott mintánál „extrémebb” (azaz bizonyos kritériumok alapján kevesebb hibát tartalmazó) lehetséges minták. Nevezzük ezeknek a lehetséges mintáknak a halmazát S -nek! Mivel S -et alapvetően befolyásolja, pontosan meg kell határozni az „extrémebb kimenetel” fogalmát. A szerzők a cikkben két kritériumot alkalmaznak, melyeknek egyszerre kell teljesülniük (az így kapott halmaz neve „step down S ”):

1. a hibás tételek száma nem haladja meg a minta hibás tételeinek számát, illetve
2. a hibák összértéke nem haladja meg a minta hibáinak összértékét.

A második lépésben meghatározzuk azon sokasági (p_i) paraméter-együtteseket, melyekre a feltételes valószínűségek összege S -halmaz felett legalább akkora, mint az előre rögzített megbízhatósági szint inverze (α) . Ezen sokasági paraméter-együttesek halmaza mint konfidenciahalmaz felett maximalizálva $\sum_{i=0}^{100} \frac{i}{100} \cdot p_i$ értéket, kapjuk meg a felső határ multinomiális becslését.

Noha ennek a becslésnek sem ismert a valódi megbízhatósági szintje, szimulációk alapján állítható, hogy nagyon közel van a névlegeshez, és emiatt a becslés sokkal pontosabb, mint akár a Stringer-képlet, akár a cellamódszer által adott becslés. Mindezen jó tulajdonságai ellenére nem annyira elterjedt, mint az előbbieket, ugyanis az 1980-as években még kevés volt a számítástechnikai kapacitás a második lépés konfidencialmazásának megalkotásához, és az azon történő optimalizálásnak a végrehajtásához.

4. Következtetések

Ma már egyre inkább ellenőrizhető, reprodukálható és módszertanilag is korrekt tevékenységet várunk el minden szakmától. Mint ahogyan az auditorok is pozitívabban ítélik meg azon szervezetek pénzügyi beszámolóit, ahol a belső folyamatok egy jól átgondolt szabályozást követnek, és nem esetlegesek, éppen így az auditori tevékenység objektivitásának növekedése is előrelépésnek tekinthető. A most bemutatott becslőfüggvények kifejlesztői úttörő szerepet játszottak a könyvvizsgálat tudományos alapokra helyezésében. A becslőfüggvényekre és a kapcsolódó mintavételi technikákra számos szoftver született (például IDEA, ACL), melyek egyre nagyobb népszerűségnek örvendenek a könyvvizsgálók között.

Amint azonban az előzőkből is kitűnik, ez az állapot sokkal inkább tekinthető egy folyamat kezdetének, mint a végének. Ezekkel a becslőfüggvényekkel kapcsolatosan még számos kritika felvethető. Csupán szimulációkkal bizonyított, hogy legalább névleges szinten megbízhatók, ami részben megkérdőjelezi alkalmazásuk korrektségét. A szimulációk alapján a módszerek túlságosan is konzervatívak, ami jelentősen csökkenti a pontosságukat és így növeli az auditált szervezet kockázatát. Alkalmazásuknak sok olyan előfeltétele van (például csak túlértékelés lehetséges), mellyel az alkalmazók nincsenek tisztában, és így – minden jószándék ellenére – téves eredményekre jutnak (különösen veszélyes ez az auditálást támogató szoftverek alkalmazásakor). További gondot okoz az eredmények helyes értelmezése.

A jövő feladata további lépések megtétele a problémák megszüntetésére, tehát olyan mintavételi eljárások és becslőfüggvények kifejlesztése, melyek egy részről bizonyíthatóan a névleges bizonyossági szintű eredményeket adják, más részről pedig univerzálisak, tehát lehetőség szerint minimális előfeltételt használnak. A könyvvizsgálók (tovább)képzése során pedig alapvetően fontos a statisztikai mintavételi módszerek, a mintavételt és a kiértékelést (becslést) támogató szoftverek hangsúlyosabb ismertetése.

Irodalom

- ARENS, A. A. – LOEBBECKE, J. K. [1997]: *Auditing: An integrated approach*. Prentice-Hall. London.
- BICKEL, P. J. [1992]: Inference and auditing: The Stringer bound. *International Statistical Review*. 60. évf. 1. sz. 197–209. old.
- Caseware IDEA Research Department [2003]: *Monetary unit sampling technical specification*. <http://www.caseware-idea.com>.
- CaseWare IDEA Research Department [2003]: *White papers on attribute sampling technical specification*. <http://www.caseware-idea.com>.
- DAVID, H. A. [1981]: *Order statistics*. Wiley. New York.
- DE JAGER, N. G. – PAP GY. – VAN ZUIJLEN, M.C.A. [1997]: Facts, phantasies and a new proposal concerning the Stringer bound. *Computers and Mathematics with Applications*. 33. évf. 10. sz. 37–54. old.
- FIENBERG, S. E. – NETER, J. – LEITCH, R. A. [1977]: Estimating the total overstatement error in accounting populations. *Journal of the American Statistical Association*. 72. évf. 295–302. old.
- GOODFELLOW, J. L. – LOEBBECKE, J. K. – NETER, J. [1974]: Some perspectives on CAV sampling plans I-II. *CA Magazine*. October, 23–30. old., November, 46–53. old.
- HALDENE, J. B. S. [1945]: On a method of estimating frequencies. *Biometrika*. 33. évf. 222–225. old.
- HAM, J. – LOSELL, D. – SMIELIAUSKAS, W. [1985]: An empirical study of error characteristics in accounting populations. *Accounting Review*. 60. évf. 387–406. old.
- HANSEN, M. H. – HURWITZ, W. N. [1943]: On the theory of sampling from finite populations. *Annual Mathematical Statistics*. 14. évf. 4. sz. 333–362. old.
- HORVITZ, D. G. – THOMPSON, D. J. [1952]: A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*. 47. évf. 12. sz. 663–685. old.
- HUNYADI L.–VITA L. [2004]: *Statisztika közgazdászoknak*. Központi Statisztikai Hivatal. Budapest.
- HUNYADI L. [2001]: *Statisztikai következtésetelmélet közgazdászoknak*. Központi Statisztikai Hivatal. Budapest.
- JOHNSON, J. R. – LEITCH, R. A. – NETER, J. [1981]: Characteristics of errors in accounts receivable and inventory audits. *Accounting Review*. 56. évf. 270–293. old.
- KAPLAN, R. S. [1973a]: Stochastic model for auditing. *Journal of Accounting Research*. 11. évf. 38–46. old.
- KAPLAN, R. S. [1973b]: Statistical sampling in auditing with auxiliary information estimators. *Journal of Accounting Research*. 11. évf. 238–258. old.
- LEHMANN, E. L. [1959]: *Testing statistical hypotheses*. Wiley. New York.
- LESLIE, D. A. – TEITLBAUM, A. D. – ANDERSON, R. J. [1980]: *Dollar-Unit Sampling-A practical guide for auditors*. Pitman. London.
- LÖLBERG T. [2004]: A sokasági arány meghatározására irányuló statisztikai eljárások véges sokaság és kis minták esetén. *Statisztikai Szemle*. 82. évf. 12. sz. 1053–1076. old.
- NETER, J. – KIM, H. S. – GRAHAM, L. E. [1984]: On combining Stringer bounds for independent monetary unit samples from several populations. *Auditing*. 4. évf. 1. sz. 74–88. old.
- NETER, J.–LOEBBECKE, J. [1975]: *Behavior of major statistical estimators in sampling accounting populations – An empirical study*. AICPA. New York.

- NETER, J. – LOEBBECKE, J. [1977] On the behavior of statistical estimators when sampling accounting populations. *Journal of the American Staistical Association*. 72. évf. 501–507. old.
- Panel on Nonstandard Mixtures of Distributions – TAMURA, H. ET AL. [1989]: Statistical models and analysis in auditing. *Statistical Science*. 4 évf. 1. sz. 2–33. old.
- PAP GY. – VAN ZUIJLEN, M. C. A. [1995]: The Stringer bound in case of uniform taintings. *Computer Mathematical Applications*. 29. évf. 10. sz. 51–59. old.
- PAP GY. – VAN ZUIJLEN, M. C. A. [1996]: On the asymptotic behaviour of the Stringer bound. *Statistica Neerlandica*. 50. évf. 3 sz. 367–389. old.
- PAP GY. – VAN ZUIJLEN, M. C. A. [2000]: Modified Stringer bounds. *Publicationes Mathematicae. Debrecen*. 57 évf. 1–2. sz. 163–183. old.
- STRINGER, K. W. [1963]: *Practical aspects of statistical sampling in auditing*. Proceedings of Business Economics, Statistics Section. American Mathematical Association. Washington. Munkanyag.
- STRINGER, K. W. [1979]: Statistical sampling in auditing. The state of art. *Annual Accounting Review*. 1. sz. 113–127. old.
- VAN HEERDEN, A. [1961]: Steekproeven als Middel van Accountantscontrole. *Maandblad voor Accountancy en Bedrijfshuishoudkunde*. 11. sz. 453. old.

Summary

This paper presents an overview of the methods used to estimate the total amount of errors in a financial report. Auditing is a special area where the application of standard estimation procedures based on well-behaved distributions can lead to inappropriate results mainly because the error distribution in financial reports cannot be considered well-behaved, showed by many empirical papers in the 1970s and 1980s. After presenting the special properties of auditing populations the paper outlines the most widely used estimators.

Közlekedési infrastruktúra és jólét Kelet-Magyarországon*

Ohnsorge-Szabó László

közgazdász-filozófus

E-mail: laszlo.szabo@pm.gov.hu

A tanulmány az autópályák regionális fejlesztési hatásait vizsgálja, alapfeltevésként fogadva el, hogy ez összefüggésben áll a régió időbeli elérhetőségével. Bemutatja és értékeli az autópálya-építések mellett érvelő hazai empirikus vizsgálatok módszerét. Rámutat arra, hogy ezek a vizsgálatok ugyan számos összefüggést tártak fel, azonban még a módszertanilag leginkább helytállóak is csak korlátozottan vehetők alapul fejlesztés- vagy regionális politikai döntésekhez, mivel az autópályák és a jólét közötti oksági összefüggések megállapításra nem alkalmasak. Ugyan ez a tanulmány sem vállalkozhat oksági összefüggések megállapítására, azt viszont bizonyítani kívánja, hogy a hazai kutatók indokolatlanul szűkítették figyelmüket kizárólag az autópályára. Ha a tágan vett közlekedési, azaz a közút- és vasút-hálózatot is tartalmazó infrastruktúrát vizsgáljuk, akkor az autópályák egyoldalú előnyben részesítését nem találjuk indokoltnak.

TÁRGYSZÓ:
Közlekedés.
Területi statisztika.
Modellépítés.

* Köszönetet mondok a tanulmány megírása során nyújtott segítségért *Scharle Ágotának, Fleischer Tamásnak, Romhányi Baláznak, Muszély Györgynek, Sáreczky Angélnak és Deli Gergelynek* (Levegő Munkacsoport).

Az írás célja az autópályák *regionális fejlesztési* hatásának vizsgálata, *abból a feltevésből kiindulva, hogy ez a hatás a régió időbeli elérhetőségével függ össze.*¹ Ha az autópályák ettől eltérő módon fejtenék ki említett hatásukat, akkor ahhoz másfajta megközelítést kellene alkalmazni. Az autópályákkal kapcsolatban felhozott egyéb pozitívumokkal (fogyasztói jólét, kényelem, biztonság) nem kívánok *részleteiben* foglalkozni. Olykor felvetik, hogy az autópályák építéskor létezne egyfajta kötelezettség a környező országok, illetve az európai hálózatok hazai területen és forrásból való kiegészítésére és továbbépítésére. Ez a probléma azonban nem képezi a tanulmány tárgyát.

1. A magyar szakirodalom

A hazai elméleti és empirikus közgazdasági irodalom az autópályák ügyében nem ellentmondásmentes. Sokszor előfordul, hogy csak általánosságok szintjén beszélnek az autópályá-fejlesztések előnyeiről, illetve ezek pozitív regionális hatásairól. Mindazonáltal születtek empirikus vizsgálatok a témában, és léteznek olyan tanulmányok is, amelyek elsősorban nem az infrastruktúrával és az autópályával foglalkoznak, de mintegy mellékesen mégis releváns tényeket és összefüggéseket tárnak fel. Az ezekből leszűrhető következtetések némiképp ellentmondásosak.

Kullman Ádám [1999] a kistérségi jövedelmeket keresztmetszeti elemzésben próbálta magyarázni a hegyeshalmi vagy a soproni határátkelőtől mért időbeli elérhetőséggel, és azzal, hogy kistérségi központ megyei jogú város-e, vagy sem. Számításai szerint 10 perces elérési időcsökkenés 3 ezer forinttal emeli meg az egy főre eső jövedelmet. Ugyanakkor felhívja a figyelmet arra, – az egyébként komoly problémára – hogy az oksági összefüggés a területi fejlődés és az infrastruktúra között statisztikai eszközökkel nem vizsgálható.² *Kullman* azt írja, hogy az idősoros vizsgálat kevésbé ígéretes, mivel Magyarországon „az elementáris erejű folyamatok közül nehezen szűrhető ki magának az infrastruktúra fejlesztésnek a hatása” (*Kullman* [1999] 19.

¹ Elérhetőségen egyesek távolsági, mások az időbeli elérhetőséget értnek. Mi időbeli, tehát a közlekedési sebességet figyelembe vevő elérhetőséget vizsgálunk.

² Erről egyébként *Kullman* [1999] is megfelelően, mikor eredményeit interpretálja. Ugyanis a keresztmetszeti modell eredménye alapján állítja azt, hogy 5 000–15 000 forint egy főre jutó jövedelemtöbblet képződik évente az érintett térségben.

old.). Mindazonáltal az oksági összefüggés teszteléséhez aligha kerülhető el idősoros vagy paneladatok használata.

A Kopint-Datorg kutatói (Bartha–Klauber [2000]) az autópálya jelenléte/távolsága szerint négyfajta kistérség-típus mutatóit vetették össze az M5-ös autópálya mentén számos indikátor alapján (munkanélküliség, népességmegtartó erő, külföldi befektetések, beruházások, export, nyereséghányad alakulása). Szerintük az előnyök egy 20-25 kilométeres sávban jelentkeznek. Bartha és Klaubert úgy fogalmaznak, hogy az autópálya csak szükséges feltétele a fejlődésnek, de nem elégséges. Másfelől viszont konkrét többletnépességet, munkahelyet, beruházást, nyereséget számszerűsítenek abban az esetben, ha Szegedig megépülne az autópálya. Azt is megemlítik, hogy miközben az autópálya menti településeken nettó népességnövekedés történt, az attól távolabb levőkön viszont hasonló mértékű a csökkenés.³

Szalkai Gábor [2001] hozzáférhetőségi viszonyokat modellező tanulmánya azért fontos, mivel az autópályákat a közlekedési infrastruktúra egészében helyezi el. Figyelemre méltó következtetése, hogy a legkedvezőbb vasúthálózati helyzetű térségek köre részben megegyezik a közútiéval, valamint, hogy a közúthálózat „teljesebb” kiépítettsége miatt a közút és a vasút versenyképességének viszonya a vasúti közlekedés térségenkénti „fejlettségétől” függ.

Az autópályák építésével kapcsolatban szkeptikusabb álláspontot foglalókat képviseli Fleischer et al. [2002] tanulmánya. Ez a műltra vonatkozóan nem talál bizonyítékot az autópályák pozitív regionális fejlesztési szerepére, másfelől nem zárja ki, hogy ilyen létezik, mivel a maga által használt módszert nem találja kellően jónak. Fleischerék tanulmánya megyei szinten végez összehasonlítást, és arra a következtetésre jut, hogy az autópályával érintett megyék nem gyarapodtak gyorsabban és a foglalkoztatottság ott nem nőtt jobban, mint a többiekben. Az írás összességében a közúti hálózat jelentőségét hangsúlyozza, amelynek csak egy eleme az autópálya.

Tanulmányunkban nem foglalkozunk azzal a kérdéssel, hogy a *tervbe* vett gyorsforgalmi útvonalak elsődlegesen hazai regionális fejlesztési vagy egyéb gazdasági célokat szolgálnak. Az előbb említett cikk szakértői későbbi írásukban azonban kitérnek erre is. Arra a következtetésre jutnak, hogy a tervek a *tranzithálózatok* fejlesztésére koncentrálnak a helyi hálózatokkal szemben, ami a koncentráció irányában hat, és ellentmondásba kerül az ország egyközpontúságának csökkentéséről szóló hivatalos nyilatkozatokkal (Fleischer et al. [2002]). Ezt a nézetet fogadja el a Levegő Munkacsoport is (Kiss–Lukács [2003] 90. old.). A gyorsforgalmi úthálózat-fejlesztési koncepció regionális fejlesztési hatása csak egy nagyon szűk körben érződik. A nagy régiók közötti kapcsolatok kiépülésében a tízéves autópálya-építési program ugyan segítségül szolgálhat, de a régióon belüli, város-falu egyenlőtlenségek kiegyenlítését nem lehet várni

³ Ez, ha nem is bizonyítékul, de indikációul szolgál arra, hogy a pálya melletti települések többleterőforrásait máshonnan szívják el; ami az autópálya iránt szkeptikusok szerint az autópálya mentén tapasztalható fellendülést magasabb, országos szinten ellensúlyozza.

tőle.⁴ Az utóbbihoz – ha közútfejlesztésben gondolkodunk – több helyütt, például Szabolcs-Szatmár-Bereg megyében az alsóbbrendű úthálózat fejlesztésére és felújítására lenne szükség (*Felischer et al.* [2002] 52. old.). A várt pozitív hatások a vasút fejlesztésével is elérhetőek lennének, sokkal olcsóbban, a lakosság szélesebb rétegei számára elérhető és inkább környezetbarát módon.⁵ A tanulmány megállapítja, hogy a hazai turizmust nem a közúti infrastruktúra hiánya fogja vissza, hiszen a világ turistaforgalmából való részesedésünk igen jelentős; hanem az, hogy a turisták csak átmennek az országon, így a turistaforgalomtól jelentősen elmarad az ebből származó bevétel (*Felischer et al.* [2002] 52. old.).

A Levegő Munkacsoport különböző tanulmányai (leginkább *Kiss–Lukács* [2003]) jellemzően olyan érveket sorakoztatnak fel, amelyek mikroszintű folyamatokra és összefüggésekre utalva mutatnak rá arra, hogy azok a mechanizmusok, amelyeken keresztül az autópályáknak hozzá kellene járulniuk a jóléthez egy adott régióban, nem működnek. (Például az autópályán való elérhetőség olyan kis időbeli megtakarítást jelent, hogy az az egyéb termelési tényezőkben mutatkozó eltérésekhez képest elhanyagolható.) Az autópálya ellen felhozott érvek egy része annak regionális fejlesztési hatékonyságát vonja kétségbe, mások az autópálya-építés negatív környezeti hatására, a fenntartható fejlődésnek ellentmondó jellegére hívják fel a figyelmet.⁶ Cáfoltják azt az érvet is, mely szerint az autópálya növeli a közlekedésbiztonságot.

Ismeretesek olyan tanulmányok, amelyek egy adott autópálya hatását a pályamenti vállalatok kikérdezésével próbálták felmérni (*Vörös–Polányiné–Czeplédi* [2003]).⁷ Az M0-ás autópálya mellé 1997 és 2000 között települt cégek tevékenysége meglehetősen visszafogottan támasztja alá az autópályához fűzött reményeket. Egyfelől a közlekedési infrastruktúra megépítése után a vállalatok évekig nem tartották a térséget befektetésre érdemesnek. A megtelepült cégek 75 százaléka nem új vállalkozás, amiből következően gyanítható, hogy mindössze tevékenység-áthelyezés történt. A cégeknek csak 28 százaléka települt az M0-ás autópálya mellé expanziója részeként, 44 százalékuk pedig a remélt magasabb profit miatt. A foglalkoztatottak száma nagyon enyhén növekedett a vizsgált időszakban, a tanulmány szerint „természetes bővülés” szerint (amit úgy értelmezzük, hogy az országos átlagnak megfelelő-

⁴ *Kullman* [1999] megjegyzi, hogy a nyertesek a jobb helyzetű rétegek lesznek.

⁵ *Felischer et al.* ([2002] 71–75. old.) a vasutat mint helyettesítőt csak a személyforgalom esetében vizsgálják, és ez alapján mutatnak rá korlátozott lehetőségeire. A teherforgalom esetében ők is a vasúti kapacitások kihasználatlanságát állapítják meg, de a forgalom vasútra tereléséhez a közúti forgalom korlátozására lenne szükség, amit nem tartanak reális lehetőségnek.

⁶ Mivel ezeket a felvetéseket máshol (*Ohnsorge-Szabó–Kajner–Ungvári* [2005] 179–202. és 270–284. old.) már összefoglaltuk, ezért az egyébként igen fontos ökológiai, gazdasági ellenvetések részletezésére nem térünk ki. Fő célunk a fejlesztéspolitika környezetileg talán kevésbé tudatos képviselőinek meggyőzése számukra is releváns érvek alapján.

⁷ Érdekes módon e tanulmány már az autóúttól 5-6 kilométerre levő vállalatok esetében is bizonytalan abban, hogy azok az autóút-beruházáshoz köthetők lennének. Ez a felfogás meglehetősen eltér az általános vélekedéstől.

en), azon belül a felsőfokú végzettségűek száma nem változott (bár 25-30 százalékuk magasan kvalifikált). A cégek nyeresége csökkent. A tanulmány készítői bizonyos fajlagosakat számolnak, így például kimutatják, hogy a 30 kilométeres M0-ás autót mentén kilométerenként 2 milliárd forint befektetésére került sor illetve hogy 600 foglalkoztatott talált állást. Ebből azonban szerintem tévesen (de legalábbis aligha kellő alappal) következtetnek arra, hogy 1 kilométer autótútnak GDP-t növelő hatása 1-2 milliárd forint évente. Ennek a következtetésnek az a megállapításuk mond el- lent, mely szerint jórészt tevékenységáthelyezés történt az ide települő cégek eseté- ben.⁸ Igen figyelemre méltó, hogy az M0-ás autópálya mellé települt cégek 24 száza- léka tervezéssel és kutatásfejlesztéssel foglalkozik, ez az ágazat pedig nem szállítás- igényes, így aligha az elérhetőség javulása volt az idetelepülés indítéka.

Vörös–Polányiné–Czeglédi [2003] M0-ás autópályára vonatkozó eredményei nem igazolják a korábban említett *Bartha–Klauber* [2000] által az M5-ös autópályáról írt tanulmány tézisét. Utóbbiak szerint a gépipar esetében döntő telepítési tényező az au- topálya. Azonban a *Vörös–Polányiné–Czeglédi* [2003] által vizsgált vállalatok szá- mában a nehéz-, élelmiszer- és könnyűiparon kívüli egyéb ipar (ahova a gépipar is tartozna) aránya mindössze 3 százalék volt. Persze az ilyen mikroszintű és a tőke, il- letve foglalkoztatás nagyságától független mutató, ráadásul két különböző autópálya esetében nem perdöntő cáfolat egyik tanulmány eredményeivel szemben sem. Arra azonban felhívhatja a figyelmet, hogy az előnyökről általánosságban beszélni prob- lematikus lehet.

Az autópályákat prioritássá tevő, határozott pozitív hatásuk mellett állást foglaló elemzések közé tartozik a GKI Gazdaságkutató Rt. két tanulmánya. Az első (*GKI* [2003a]) tanulmány az M3-as és az M5-ös autópálya által érintett kistérségeket ha- sonlítja össze. Ezeket három csoportba sorolja: 1. az autópálya által közvetlenül érin- tettek, 2. közvetve érintettek (20 perenyire vannak attól), és 3. a kontrollcsoportként használt többiek csoportjára. Azt találja, hogy az első két kistérségtípusban 1992 és 2001 között számottevően nagyobb mértékben nőtt az exportárbevétel, a külföldi tő- kebefektetés, nőttek a bérek, nőtt az szja és a társasági nyereségadó (tánya-) bevétel, és valamelyest jobban csökkent a munkanélküliség, mint a kontrollcsoportban.⁹

A másik tanulmány (*GKI* [2003b] ennek rövidített változata: *Bíró–Molnár* [2004]) hasonló módszert választott az autópályák hatásának vizsgálatára, amennyi- ben azt vizsgálták, hogy az adott kistérség „beleesik-e” a „használatba vett autópá- lyák 10, illetve 30 kilométeres vonzáskörzetébe, vagy sem” (*Bíró–Molnár* [2004]

⁸ Mindazonáltal megjegyzik, a foglalkoztatási hatás más autópályák esetén az M0-ás autópályánál mért foglalkoztatási fajlagosnak csak harmada-negyede lehet. Persze ez a szám is csak annyira megalapozott, mint a fajlagos alapul vétele a foglalkoztatási hatás közelítéseként.

⁹ Némiképp problémát jelent, hogy nem magyarázzák meg, mit értenek egy kistérség autópályától való távolsá- gán (legközelebbi pontjának vagy településének távolsága, központjának távolsága, esetleg településeinek átlá- gos vagy súlyozott távolsága stb.).

1060. old.). Az autópálya-építéssel kapcsolatos következtetések egyértelműek. Egy helyen, az észak-alföldi régióval kapcsolatban azt írják, hogy csak az *infrastrukturális* fejlesztések (*gyorsforgalmi* utak) hozhatnak fellendülést, más helyen pedig, hogy a „lassan bővülő hazai autópályák gazdaságélénkítő hatása nem kérdéses” (1060. old.). A *szélesebben* felfogott infrastruktúra hatásával kapcsolatos megállapításuk árnyaltabb: annak hatása a komplex fejlettségre „nem elhanyagolható, de semmiképpen sem definitív” (1056. old.). A szerzők meglegszenek – és a módszerük alapján meg is kell elégedjenek – azzal, hogy pozitív vagy negatív erős, közepesen erős és laza kapcsolatokat állapítanak meg az autópálya és az egyes faktorok között.

Igen komoly elemzést végzett a témában *Németh Nándor* [2004].¹⁰ A tanulmány azért érdekes, mivel egyfelől azt állítja, hogy az autópálya területi fejlődést generáló hatása „már évtizedek óta nem szorul bizonyításra” (141. oldal), másfelől viszont empirikus eredményei és ezzel kapcsolatban tett értékelései éppen azt mutatják, hogy meglehetősen ellentmondásos kép rajzolódik ki Magyarországon ezzel a hatással kapcsolatban.¹¹ A tanulmány felveti azt a szempontot, hogy az autópályák melletti területek jobb általános helyzete egyszerűen az érintett városok egyébként is jobb gazdasági pozíciójának tudható be. „E tény erősen megnehezíti az autópályák önálló területfejlesztő hatásának kimutatását” írja (*Németh* [2004] 149. old.). *Németh* végkövetkeztetése az, hogy igen óvatosan kell kezelni az autópályák közvetlen területfejlesztő hatását, és ennek a hatásnak a kibontakozásához egyéb dinamizáló feltételek szükségesek (177. old.).¹²

Németh, sokakkal együtt, az elérhetőség javításában látja az autópályák pozitív szerepét. A probléma megítélésem szerint azonban ezzel a tanulmánnyal is az, hogy az elérhetőség javítását azonosítja az autópálya-építéssel. A magam tanulmánya abban különbözik *Némethétől*, hogy az elérhetőség javításának lehetőségeit a szóba jöhető három mechanizmuson keresztül veszem számításba (autópálya-építés, egyéb közúti és vasúti fejlesztések), nem pedig egyedül az autópálya-építésen keresztül. Et-

¹⁰ Bizonyos értelemben az ő módszertana és modellje áll a legközelebb az én tanulmányoméhoz. Települési szintű adatokat (is) elemez, a jövedelmeket és a munkanélküliséget tekinti függő változónak, a magyarázó változók között a képzettségi szint, a nyugati határtól való távolság és az autópályától való távolság szerepel.

¹¹ Például egy helyütt úgy fogalmaz, „nem kétséges, hogy Somogy számára mekkora jelentősége van a sztrádának”, aztán megjegyzi, hogy az M7-es autópályával kapcsolatban a helyiek egy része attól fél, hogy a potenciális balatoni vendégek jelentős része az autópálya miatt az Adriát fogja választani. (*Németh* [2004] 150. old.)

¹² A tanulmány értéke, hogy időbeni összehasonlításokat is végez (ezek nem időszerelemzések). Ennek során azt vizsgálja, hogy egy új autópálya-szakasz (az M1-es és az M3-as autópályák esetében) átadása révén egy 50 kilométer átmérőjű körön belülré került településeken a cégsűrűség hogyan alakult azokkal a településekkel összehasonlítva, amelyek már az átadás előtt is az 50 kilométeres sávon belül voltak. (*Németh* [2004] 174-176. old.) Az M1-es autópálya esetében az átadás nem jár a cégalapítás felgyorsulásával az autópálya 50 kilométeres sávján belülré kerülteknél, az M3-as esetében viszont igen. Utóbbinál viszont azt látjuk (amire *Németh* nem hívja fel a figyelmet), hogy a korábban is a sávon belül levőknél hasonló ütemben fokozódott a cégalapítási láz. A cégalapítás következtében egyébként, mint *Németh* maga jegyzi meg, az M3-as autópályán a foglalkoztatás és a jövedelmek bővülése nem volt tapasztalható.

től függetlenül – mint látni fogjuk – eredményeim sok szempontból harmonizálnak Némethével, csak éppen nem használhatók ki még olyan mértékben sem az autópályák javára, mint az övéi.

Németh Nándor számítás szerint az autópálya-távolság semmilyen szerepet nem játszik a külföldi tőke térnyerésében (Németh [2005] 171. old.).¹³ A külföldi tulajdonban levő vállalatok székhelye 60 százalékban Budapesten van, az autópálya 50 kilométeres körzetében csak 17 százalékuknak található.

Tóth Géza [2005] az autópályától várható potenciálváltozás hatását számszerűsíti. Egy település potenciálja azt fejezi ki, hogy egy adott település mennyire van közel tehető településekhez. A potenciál annál nagyobb, minél nagyobb GDP-jű települések találhatóak minél közelebb a szóban forgó településhez. Az autópálya arra képes, hogy időben közelebb hozza a többi települést, és ezzel növelje a települések potenciálját. Az ezzel a módszerrel kapcsolatos kifogást maga a szerző említi: ha az autópályával a potenciálnövekedést akarjuk maximalizálni, akkor minél inkább a fejlett településekhez közel kell telepíteni azt. Továbbá a módszer feltételezi, hogy az építés hatására kialakuló fejlettségi növekmény a jelenlegi fejlettségi szinttel egyenesen arányos. A szerző szerint az autópályától azt lehet várni, hogy a megyék közötti egyenlőtlenségeket fogja mérsékelni, a megyén belüli egyenlőtlenségek viszont nőni fognak. A periferikus településeken – ha közlekedési infrastruktúrában gondolkodunk – az alsóbbrendű utak fejlesztésével lehet segíteni (Tóth [2005] 10. old.). Utóbbi megállapítása bizonyítja, hogy a szélesebb közlekedési infrastruktúra fontosságát nem lehet tagadni.

A nem kimondottan az infrastrukturális fejlesztésekkel foglalkozó, ám témánk szempontjából érdekes tanulmányok (Kertesi [2000b], Köllő [1997]) az autópályától várt kedvező hatás két főbb elemével, a munka és a tőke mobilitásával kapcsolatban szolgáltatnak információt. Egyik feltevés szerint a tőke könnyebben „mozogna” az autópályákon a fejletlenebb régiók felé, a másik szerint a munkavállalók tudnának könnyebben eljutni a kedvezőbb helyzetű régiókba. Mindkét logika szerint tulajdonképpen a két fő termelési tényező kerülne közelebb egymáshoz, javulna a kölcsönös elérhetőségük.

Előzetesen csak annyit, hogy az 1990-es népszámlálás alapján tudjuk (az újabb népszámlálásban nincsenek már ilyen adatok), hogy az ingázók célpontja többnyire magában a megyében található.¹⁴ Ha megnézzük egy adott megyében azokat a településeket, amelyek nettó fogadók, és amelyek nettó kibocsátók, akkor azt látjuk, hogy az ezek közötti út-távolságokban az autópályák igen kis arányt tesznek ki, ha egyáltalán szerephez jutnak. Persze nem kizárt, hogy 1990 óta megugrott a megyén kívülre ingázók aránya, így már nem releváns az 1990-es statisztika. Azonban a

¹³ A külföldi tőke lakosságarányos nagysága csak egyes években és gyengén szignifikáns.

¹⁴ Borsod-Abaúj-Zemplén megyében 97 százalék a KSH [2004] 1990. évi népszámlálási területi adatainak alapján.

Kertesi [2000] által hivatkozott 1996-os mikrocenzus – mely az 1990-es évek egyetlen, az ingázásra vonatkozóan kellően informatív felmérése – szerint a képlet alapvetően nem változott: az ingázók zöme, 71 százaléka falusi lakos, és 56 százaléka faluból városba ingázik.

Köllő [1997] és *Kertesi* [2000b] elemzése szerint elsősorban a helyi érdekű tömegközlekedés (vasút és Volán-buszok) menetrendjének foglalkoztatásbarát átalakításával lehetne jelentősen csökkenteni legalább a regionális munkanélküliségi rátában mutatkozó különbségeket, de akár a munkanélküliség szintjét is. *Kertesi* [2000b] azt a kérdést is felveti, nem kellene-e a közlekedés és a regionális politika prioritásait újragondolni, és a hangsúlyt a kistérségen *belüli* közlekedés fejlesztésére, és a távolsági buszjáratok támogatására áthelyezni. 1990-es évek közepi adatokra építve az említett két szerző rámutat, hogy a béregységben mért autózási költség sokszorosa a nyugat-európainak, így nálunk a tömegközlekedés sokkal erősebben befolyásolhatja a munkaerő-piaci egyenlőtlenségeket. Az 1990-es évek közepe óta nyilván növekedtek a jövedelmek, – és az autózást terhelő adók *relatív*e csökkentek – ám ennek ellenére is érdemes utalni arra a megállapításra, mely szerint a települések több mint felében a munkanélküliségtől fenyegetett népesség átlagfizetésének fele sem volt elegendő a második legközelebbi város naponkénti megközelítésére.¹⁵ 10 kilométeres ingázás esetén az állások egyharmada nem térítene meg a minimálbér és az autós ingázás költségei közötti különbséget, tömegközlekedési kapcsolat esetén viszont ez az arány csak 15-20 százalék. (20 kilométeres ingázásnál ezek az arányok kétharmad, illetve 30-35 százalék).¹⁶ 1993-as árakon 2000-6000 forint közöttire teszik azt a havi, személyenkénti utazási pluszköltséget, amely a munkaerőpiac kitérítésében szerepet játszhat. Az 1990-es évtized utolsó harmadában országosan elinduló kedvező foglalkoztatási folyamatok nem jelentkeztek a falvak egy jelentős részében (*Kertesi* [2000b]). Az évtized végén tapasztalható konjunktúrából való kimaradást – *Kertesi* [2000b] szerint – az ingázás költségeinek továbbra is prohibítív szintjével lehet indokolni; azt nem magyarázhatta a képzetesebb munkaerő elvándorlása, mivel annak mértéke igen csekély.

A munkagazdászok kimutatták, hogy a tömegközlekedési kapcsolat sűrűsége – a mérvadó, az említett 2000-6000 forintos költségen elérhető távolságban nulla, illetve négy települési kapcsolattal rendelkező települések között – 5-6 százalékpontos különbséget eredményez azonos kistérségen belül is a munkanélküliségi rátában. A viszonylag sűrű településhálózat és szerény távolságok ellenére kevés központ érhető el vonattal, busszal. A települések 60 százaléka esetében 40 kilométeres körzetben található minden 3-4 kistérségi központ a településeknek csak 1 százalékából érhető

¹⁵ A szerzők nem azt az utat választják, hogy az autózást terhelő elvonások csökkentése mellett kardoskodjanak, amit mi sem tartunk járható útnak. Bár a szokásos panasz a rendkívül magas magyarországi üzemanyagárak miatt itt is megjelenik.

¹⁶ *Köllő* [1997] a munkanélküliek korábbi keresetével is összehasonlítva az ingázás költségeit megállapítja, hogy 20 kilométeres ingázás a kereset 40 százalékát viszi el, 40 kilométeres pedig 80 százalékát.

el a reggeli órákban. Tíz munkanélküliből négynek nem volt esélye saját kistérségén kívülre ingázni.

Köllő [1997] 500 olyan településről beszél, ahol a tömegközlekedés milyensége jelenti a fő mobilitási korlátot, azaz ahol a magas munkanélküliségtől sújtott településeket jó állapotú városi munkaerőpiacok veszik körül, viszont a munkavállalók oda például a menetrendnek a munkakezdéshez nem igazodó jellege miatt nem tudnak bejutni.¹⁷ *Kertesi* [2000b] elemzésének fontos tanulsága, hogy a szorosan vett lakóhely hatása az ingázási/munkához jutási esélyekre kétszer erősebb, mint a regionális különbségeké. Ez igencsak indokoltá teszi a közlekedési infrastruktúra-fejlesztések esetében a települési szintű megközelítés alkalmazását; a regionális, kistérségi szintű elemzések esetleg túlzottan elnagyoltnak bizonyulhatnak.

A munkaerőnek nemcsak napi vonatkozásban értelmezhető a mobilitása, hanem településváltásként is. *Kertesi–Köllő* [1998] megállapítják, hogy a magyar népesség mobilitása egyáltalán nem jelentéktelen, évente 3,5-4,6 százalékra tehető a települést váltó népesség aránya. Ebben a fajta mobilizációban az autópályák léte vagy nem léte, sőt általában a közlekedési rendszer milyensége egy bizonyos – nálunk minden kétséget kizáróan meglevő – úthálózati minőség felett nem sokat számít.

Az autópályák és a tőke mobilitása közötti összefüggésre vonatkozóan kevesebb következtetést tudunk levonni. Az persze szembetűnő (még az átlagember számára is), hogy az autópályák mellett, illetve néhány csomóponton gombamód szaporodnak az üzletek. Ám megfogalmazódnak olyan fajta kételyek, hogy itt mindössze a tőke térségen, kistérségen belüli átrendeződéséről van szó; azaz amennyivel nő a tőkeállomány közvetlenül az autópálya mentén, annyival csökken attól valamelyest távolabb.

Néhány közvetett bizonyíték, illetve tapasztalati összefüggés alapján mindenesetre adódik egy-két tanulság. Tiszta versenymodellben (*Kertesi–Köllő* [1998]) mind a tőke, mind a munka olyan fajta kiegyenlítő mozgását várnánk, amely a tőkét az olcsóbb munka, a munkát a foglalkoztatást biztosító tőke felé tereli.¹⁸ Ezzel szemben a szerzők megállapítják, hogy versenymodell alapján várttal ellentétben a regionális különbségek még fokozódtak is. A magas munkanélküliségű régiókban a nyereséges vállalatok aránya már 1994–1995-ben elérte az országos átlagot, átlagos profitja pedig meg is haladta azt.¹⁹ A profítképződés, valamint az, hogy nagymértékű jövede-

¹⁷ Más kezelést kell viszont alkalmazni ott, ahol a városi központok maguk is válságban vannak.

¹⁸ *Kertesi–Köllő* [1998] szerint ezek a mechanizmusok valóban léteznek, és csak ezektől várhatjuk a regionális különbségek mérséklését. A szerzőknek ezzel a tézisével, mármint, hogy csak ezen piaci kiegyenlítő mechanizmusoktól várhatunk javulást, megítélésünk szerint nem teljesen egyezik az a más helyen kifejtett álláspontjuk, mely szerint állami eszközökkel kellene a regionális tömegközlekedést fejleszteni.

¹⁹ Témánk szempontjából fontos *Kertesi* [2000b] tanulmányának az a megállapítása, hogy a minden bizonnyal a kedvezőtlenebb foglalkoztatási helyzetű településeken élők termelékenysége szisztematikusan alacsonyabb, mint a többieké (a munkahely-találási esélyek nemcsak a céltelepülés, hanem a lakóhelyül szolgáló település foglalkoztatási helyzetétől is függnnek, még hozzá ott rosszabbak, ahol a lakóhelyen rosszabb a foglalkoztatási helyzet); illetve a munkaadók hasznot húznak a magas munkanélküliségű térségből érkezők helyzetéből.

lemátcsoportosítás ment végbe a dolgozók rovására, a munkáltatók javára, mégsem bizonyult elégségesnek ahhoz, hogy a helyi tőkeakkumulációt felgyorsítsa, befektetőket vonzzon, és kompenzálja az egyéb természetű hátrányokat. A dolgozókkal szembeni erőfölény, az alacsony bérek és a munkaerőbőség odavonozhatná a külső befektetőket, és a helyi pénztulajdonosokat is arra ösztönözhetné, hogy banki befektetés, pazarló fogyasztás helyett vállalkozásba fogjanak. A magántőke-felhalmozás feltételei adottak az elmaradott régiókban; ez megkérdőjelezi azt a hipotézist, hogy amire szükségük van az elmaradott régióknak, az a befektetést kereső tőke.

Ezen túl kétséges, hogy az – egyébként nem is annyira szűkösen rendelkezésre álló tőkét – autópályáktól remélhetnénk. *Kertesi-Köllő* [1998] megállapítása szerint összességében hiábavaló állami támogatásoktól várni az elmaradott térségek kilábalását, ha egyszer az azonos munkáért itt fizetett 15-20 százalékkal alacsonyabb bér sem volt elegendő ahhoz, hogy a régiók magukhoz csalogassák a beruházókat. Ennek ellenére az iskolázottság és a közlekedési kapcsolatok javítását az állami szerepvállalás lehetséges terepeként nevesítik. Közlekedési kapcsolatokon viszont – mint láttuk – nem az automobilizáció támogatását és az autópálya-építést értik.

Mivel a külföldi tulajdonú vállalatnál foglalkoztatottak aránya *Kertesi-Köllő* [1998] szerint igen jó közelítője az új tőkebefektetéseknek, a külföldi tulajdonú vállalatnál foglalkoztatottak arányát bontották fel 1993–1995-ben az iskolázottság és a Hegyeshalomtól való távolság függvényében. A nagy és a kis munkanélküliségű települések közötti különbséget 1995-ben 60 százalékból az iskolázottság és 30 százalékból a Hegyeshalomtól való távolság magyarázta. 1993-ban viszont még nagyjából hasonló, 40-40 százalékos volt a két tényező magyarázó ereje; azaz a humán tőke jelentősége egyre nő a geográfiai – közlekedési infrastruktúra-befektetésekkel esetleg ellensúlyozható – elhelyezkedéshez képest. A szerzők ezzel kapcsolatban felhívják a figyelmet arra, hogy a nem ricardói, hanem a mind finomabb specializációra épülő munkamegosztás korában például Fejér megyének a kapcsolatai erősebbek Komárom városával és Bajorországgal, mint Nógrád vagy Borsod megyével.

2. A módszer

Az általunk követett módszertan figyelembe veszi az említett tanulmányok eredményeit, az alkalmazott módszerek erősségeit és hátrányait. Települési szinten vizsgálódunk, és igyekszünk települési szintű adatokkal magyarázatot adni az egyes települések közötti *foglalkoztatottsági és jövedelmi* különbségekre.²⁰ Tehát csak két mu-

²⁰ Az általunk használt települési adatok forrása a KSH megyei statisztikai és az 1990-es, 2001-es népszámlálás adatai (*KSH* [2004]), valamint az APEH Adóügyi Főosztálya által készített, és részünkre megküldött 1993. és 2001. évi szja-bevallás kiemelt adatai településenként táblák.

tatót vizsgálunk – foglalkoztatás és jövedelem –, eltérően a *GKI*- [2003a és b] tanulmányoktól, amelyek több indikátort használnak. Sokkal több mutató használata megítélésünk szerint érdemben nem járulhat hozzá a tényleges jóléti hatás becsléséhez. (Például az APEH egyik munkatársa arra hívta fel figyelmünket, hogy a tása-adatok települési szinten nem mondanak sokat, mivel a telephely és az üzem egymástól igen távol is lehet, a jövedelem az üzemből képződik, viszont a telephely fizeti be az adót.) A vizsgált autópályák az M3-as és az M5-ös. Az autópályák által érintett négy megyét (Bács-Kiskun, Csongrád, Heves és Borsod-Abaúj-Zemplén), pontosabban ezek bizonyos népességszám feletti települését (Bács-Kiskunban, Csongrádban és Hevesben a körülbelül 3000 fő fölötti, Borsod-Abaúj-Zemplén megyében pedig a 2000 fő feletti népességűeket) vettük be az összesen 124 elemű mintába. A vizsgált településeken élők száma közel 1,5 millió. A teljes mintán belül megkülönböztettük a vasúti kapcsolattal rendelkező 93 település részmintáját.

3. Modellszámítások

Megállapításaink alátámasztására a következő egyszerű modellek becslésére vállalkoztunk:

$$L = a_1K_p + a_2K_h + a_3K_i + a_4K_0,$$

illetve

$$W = a_1K_p + a_2K_h + a_3K_i + a_4K_0$$

Az L a települési foglalkoztatottságot, W a települési átlagjövedelmet mutatja, K_p vektor a magántőke nagyságát hivatott kifejezni (pontosabban arra utaló proxyk szerepelnek benne), K_h a humántőke nagyságát, K_i az infrastrukturális tényezőket, köztük az érdeklődés középpontjában álló közlekedési infrastruktúrát, K_0 pedig az egyéb szerkezeti tényezőket.

A foglalkoztatottsági mutató az adott településen élők közül foglalkoztatottak számát osztja az aktívák számával. A jövedelmi mutató az APEH-től kapott kimutatás szerint a településen bevallott összes adóalapot a bevallók számával osztja. A humán tőkét olyan változókkal közelítettem, mint az egyetemi végzettségűek, valamint a nyolc osztálynál többet nem végzettek aránya, a felekezeti és nemzetiségi hovatartozást kifejező kulturális jellemzőkre vonatkozó arányok. A gazdasági tőkét a vállalkozások sűrűségével, a termelőszövetkezet létevel az adott településen, illetve a részvénytársaságok számának a foglalkoztatottakhoz viszonyított arányával gondoltam kifejezni. Az infrastrukturális tőkére utal a telefonellátottság sűrűsége, a vasúti összeköttetés léte, az au-

tópálya léte (három különböző, 20, 40 és 80 kilométeres) távolságban, a Budapest vasúti megközelítéséhez szükséges minimális idő, a Budapesttől való közúti távolság, az útszakaszok minőségével súlyozva, valamint a távolsági autóbusz-állomás léte és végül a foglalkoztatottakra eső személygépkocsik száma a településen. Egyéb szerkezeti tényezőkre utal a település városi/falusi jellegét kifejező változó, az eljárók aránya a településen munkát találó helybeliek arányában, bizonyos idegenforgalmi adottságok, az iparban és szolgáltatásokban foglalkoztatottak aránya a településen élő és ott foglalkoztatottak között és a kistérségi foglalkoztatottság aránya.²¹

A modell hasonló a *Jacoby* [2000] által alkalmazotthoz (*Onsorge-Szabó* [2005]). A bizonyos tényezők szempontjából hátrányos helyzetű települések hátrányai halmozódhatnak, mint *Kertesi* [2000b] tanulmányában láttuk. Ugyanakkor, mint *Jacoby* [2000] tanulmányából kiderült, a lakosok alkalmazkodhatnak a hátrányos helyzethez, olyan stratégiákat kialakítva, amelyekkel ellensúlyozhatják egy bizonyos szempontból adódó hátrányukat. Ezért indokolt a vizsgálatban a kvadratikus összefüggések, kvadratikus függvényforma alkalmazása. *Kertesi-Köllő* [1998] egyébként nemlineáris formát használ, amikor a gazdasági szervezetek sűrűségének szóródását az iskolai végzettséggel magyarázza.

Meg kell jegyezni, hogy a szóba jöhető változók köre települési szinten korlátozott, kissé túlértékeli lehetőségeinket ezek alapján „tőkéről” beszélni, amennyiben például a népességhez viszonyított vállalkozások számát tekintjük a magántőke „nagyságának”.²²

Az előző egyszerű egyenlet alapján többféle modell alakítható ki. Az összehasonlítás végett érdekes lehet pusztán az infrastrukturális változókkal, egyváltozós modellekkel magyarázni a foglalkoztatást és a jövedelmet, azonban az egyváltozós modellek bemutatásától ebben a tanulmányban eltekintek.²³ A többváltozós modellek között az egyik típus felülről építkezik, felhasználva az összes változót. A másik alulról építkezik, és csak a „legjobb” változókat használja. A felülről építkező modelleken belül megkülönböztettem két típust (A és B), ahol az A-modellben szerepel az összes változó, a B-ből viszont az A-modellből kiindulva a rosszul teljesítő változókat kihagytam.

3.1. Többváltozós modellek

Mind a foglalkoztatás, mind a jövedelem települések közötti különbségeinek magyarázatában többféle modellt alakítottunk ki. Egyfelől a szóba jöhető összes válto-

²¹ A használt változókat, azok definícióját és elméleti indokoltságát a *Statisztikai Szemle* interneten megjelenő számában található melléklet tartalmazza, ugyanúgy, mint a regressziók eredményét bemutató táblázatokat, mellékleteket.

²² A változók részletes bemutatásra lásd az internetes változatot.

²³ Az internetes változatban ezek megtekinthetők.

zót igyekeztünk beépíteni (felülről építkezve) a modellbe, másfelől csak a legerősebb változók szűk körét alkalmazva végeztünk számítást (alulról építkező modell). Az alulról és a felülről építkező modellek így néhány magyarázó változójukban természetesen eltérnek.

A felülről építkező modell

Első lépésben beépítettük az A-modellbe az általunk kiválasztott összes változót. Fontos, hogy a magyarázó változók között szerepelt a jövedelmi mutató is. A második, B-modell úgy jött létre, hogy az A-modellben 0,5 feletti p -értékű változók csoportját kihagytuk, majd ezután folyamatosan hagytuk el a legrosszabb p -értékű változókat, amíg minden benmaradó változó p -értéke 0,1 alá nem került, illetve amikor további inszignifikáns változó elhagyása már nem járt a modell javulásával (a Gretl ökonometriai szoftver által használt információs kritériumok többségében).

Az alulról építkező modell

Ebben az alulról felfelé építkező, minimális számú változót használó modellben a különböző közlekedési infrastruktúra-változókra külön almodelleket építettünk, és az almodellekben a közlekedési infrastruktúra-változót egészítettük ki néhány, a foglalkoztatással (jövedelemmel) igen erősen korreláló változóval; olyanokkal, amelyek elméleti erőssége ezt indokolta, és az endogenitás a függő változóval is viszonylag gyengének feltételezhető. A változók között egy-egy humán-, gazdasági-, valamint infrastruktúra-tőkét közelítő szerepel. A kiegészítés menete egyszerűen az volt, hogy az utóbbi változókkal kiegészítettük a modellt, és ha ezek szignifikánsnak bizonyultak, akkor maradhattak benne. Ez a megközelítés azzal az előnnyel kecsegtet, hogy az egymással és a függő változóval esetleg endogén viszonyban álló magyarázó változókat kihagyva, az OLS-becslés használata a korábbinál meggyőzőbb lehet.

3.2. Eredmények

A következő részben ismertetjük a különböző modellek használatával elért eredményeinket.

Foglalkoztatásmodellek

Felülről építkező modellek (1. internetes melléklet). A B-típusú, a rossz a t -statisztikájú változókat elhagyó modellekben a teljes és rész minta esetében szignifikáns változók köre némileg eltér. A rész minta alapján kapott modell illeszkedése

(szabadságfokkal korrigált $R^2=0,77$) jobb, mint a teljes mintásé (szabadságfokkal korrigált $R^2=0,65$).²⁴ A teljes mintában a minőséggel súlyozott közúti távolság volt a legerősebb közlekedési infrastrukturális mutató, míg a vasúttal rendelkező települések esetében a vasúti távolság.

Az alacsony iskolai végzettség, a jó jövedelmi helyzet, a vállalkozások sűrűsége, a településről dolgozni eljárók aránya, a kistérségi foglalkoztatás aránya a várt módon befolyásolja a foglalkoztatást. A két mintában különböző nem közlekedési infrastrukturális mutatók bizonyultak szignifikánsnak, ezek előjele is a vártak megfelelő. Úgy tűnik, hogy a várakozásoknak megfelelően mind a három vallási felekezet esetében alapunk van feltenni, hogy jelenlétük javítja a foglalkoztatottságot, azon belül a legkisebb, átlagosan kevesebb mint 2 százalékos arányt kitevő evangélikusoké a római katolikusokénál számottevően nagyobb mértékben.

Az egyéb változók között meglepetés, hogy a szolgáltató és az ipari szektorban dolgozók aránya milyen jelentős mértékben gyakorol negatív befolyást a foglalkoztatásra. Meglepő a rész minta esetében a roma népesség arányának pozitív hozzájárulása a foglalkoztatáshoz, és az is, hogy az egyetemi végzettség nem javítja – sőt a rész minta esetében még rontja is – a foglalkoztatottságot. Meglepő hogy a termelőszövetkezetek hatása vagy nem kimutatható (teljes minta) vagy negatív (rész minta). Az utóbbi nehezen magyarázható jelenségek miatt is célszerű más megközelítésben is vizsgálni az adatokat.

Alulról építkező modell (2. internetes melléklet). Az alacsony iskolázottság foglalkoztatottságra gyakorolt hatása, hasonlóan a felülről építkező modellhez negatív. A vállalkozói sűrűség és a más településre ingázók aránya ugyanúgy kedvezően befolyásolja a foglalkoztatottságot, mint a felülről építkező modellben. Meglepő az egyetemi végzettség szignifikánsan negatív együtthatója is mind az öt, különböző közlekedési infrastruktúra-változót tartalmazó almodellben, és mind a két (a teljes és a rész-) mintában.

A közlekedési infrastruktúra-mutatók közül az autópályától 80 kilométernél nagyobb távolság a szokásos szignifikanciaszintek mellett nem jó magyarázó változó, ami még inkább igaz a vasúti kapcsolat létrejöttére. A legnagyobb magyarázó erővel a közúti távolság mutatója bír. (Ha úgy indítjuk a modellt, hogy az összes közlekedési infrastruktúra-változót behelyezzük, akkor a többi közlekedési infrastruktúra-mutató inszignifikánsnak bizonyul a szelekció során.) A vasúti kontaktussal rendelkező települések esetében ugyanaz a helyzet a vasúti távolsággal, mint a teljes minta esetében a közúti távolsággal. Ennek az almodellnek a legnagyobb az R^2 értéke az alulról építkező modellek között.

A közlekedési infrastruktúra-mutatókra vonatkozó eredmények robusztusak, amennyiben a felülről és az alulról építkező modellek egyaránt azt az eredményt ad-

²⁴ Megjegyzendő, hogy a hozzánk viszonylag hasonló modellt használó Németh [2004] is 0,60-0,70 körüli R^2 értékeket kapott. (154. old.)

ják, hogy a teljes mintában a közúti, a vasúti kapcsolattal rendelkező részmintában a vasúti távolság a legjelentősebb tényező. A közúti távolság együttthatója a kapott eredmények alapján (-0,18)-(-0,19)-re tehető, a vasútié pedig (-0,14)-(-,20). Ezt azt jelenti, – a négyzetes tagot is figyelembe véve – hogy a foglalkoztatottsági hatás a közúti és a vasúti távolság átlagos (nagyjából a mediánnak is megfelelő 112 kilométeres és 142 perces) értéke alapján a felülről építkező modellben körülbelül 10-11 százalékpont, az alulról építkezőben pedig 12-15 százalékpont. Ezek az együttthatók magasabbak, mint az egyváltozós modelleknél számoltak. Figyelemre méltó, hogy a többi változó hatásának figyelembevétele nemhogy csökkentette, hanem növelte az együtttható értékét.

Jövedelmi modellek

Felülről építkező modellek (3. internetes melléklet). A teljes és részminta esetében szignifikáns változók köre némileg itt is eltér. A részminta alapján kapott modell illeszkedése (szabadságfokkal korigált $R^2=0,83$) jobb, mint a teljes mintásé (szabadságfokkal korigált $R^2=0,78$). A jövedelmek vonatkozásában, eltérően a foglalkoztatottságtól, mind a rész-, mind a teljes mintában az autópályától vett 80 kilométeres távolság bizonyult a legerősebb közlekedési változónak, a várt előjellel; de a teljes mintában – ahol ennek a mutatónak volt értelme – a vasúti pályaudvar léte is szignifikánsnak bizonyult.

Az alacsony és a magas iskolai végzettség, a foglalkoztatottság és az iparban, a szolgáltatásban foglalkoztatottak aránya a várt módon befolyásolják a jövedelmi viszonyokat. A két mintában a nem közlekedési infrastrukturális mutatók nem bizonyultak szignifikánsnak. A településről dolgozni eljárók aránya (váratlanul nem mondható módon) negatívan hat a jövedelemre. A római katolikus és evangélikus népesség esetében van okunk legalább az egyik minta alapján alacsonyabb, a németek esetében magasabb jövedelmekre következtetni.

Meglepetést a foglalkoztatottsági szempontból jelentős pozitív hatással bíró vállalkozási sűrűség erősen szignifikáns negatív együttthatója okoz. A roma népesség aránya és a részvénytársaság léte inszignifikáns.

Alulról építkező modell (4. internetes melléklet). Az alulról építkező modellben a változók előjelei – hasonlóan a felülről építkező modellhez – a vártak megfelelnek, kivéve a vállalkozási sűrűséget. A különböző, egyaránt szignifikáns közlekedési infrastruktúra-mutatók magyarázóerejében nincs nagy különbség a teljes mintában, itt is az autópályától való 80 kilométeres távolság a legerősebb. A részmintában a szokásos szintek mellett egyedül a 80 kilométeres autópálya-távolság bizonyult szignifikánsnak, a többi inszignifikáns a szokásos szintek mellett. (Ugyanakkor a közúti távolság kivételével a többi változó p -értéke is a 10-17 százalékos sávban van.) Figyelemre méltó, hogy a vasúti kapcsolattal rendelkező településeken az egyébként erős magyarázóerő-

vel bíró autópálya és közúti távolság mutatók milyen gyengén „szerepelnek”. Ez arra utalhat, hogy ahol mindkét közlekedési infrastruktúra (a vasút és a közút) rendelkezésre áll, ott a közúti elérhetőség jelentősége valamelyest csökken, a vasút képes biztosítani a helyi gazdaság fejlődéséhez szükséges hálózati kapcsolatot (ez persze nem feltétlenül jelenti azt, hogy a közúti összeköttetésnek nincs is jelentősége ezeken a településeken). Azaz nem az autópályáknak van egyedülálló jelentőségük, hanem a közlekedési elérhetőség az, ami sokkal inkább számít, és amit a vasúttal is biztosítani lehet.

Fontos kitérni arra, hogy a közlekedési infrastruktúra-mutatók együtthatói jelentősen nagyobbak, ha a foglalkoztatottsági mutató nem szerepel a magyarázó változók között. A vasúti kapcsolattal rendelkező települések mintájában ekkor az összes közlekedési infrastruktúra-mutató szignifikáns. Ennek a jelenségnek az az oka, hogy a közlekedési infrastruktúra a foglalkoztatásra gyakorolt pozitív hatásán keresztül is hozzájárul a magasabb jövedelmekhez. Azaz pozitív jövedelmi hatásukat részben a foglalkoztatottság együtthatója tartalmazza.

A közlekedési infrastruktúrára vonatkozó eredmények itt is robusztusak. A vasúti kapcsolat léte és az autópálya 80 kilométeres távolságon belüli elérhetősége egyaránt 4 százalékos mértékben növeli az átlagos jövedelmeket. A Budapesttől átlagos, (minőséggel súlyozott közúton) 112 kilométerre levő település 5 százalékos jövedelmhátrányban van a közelebb fekvőkhöz képest. Ha eltekintünk attól, hogy a vasúti távolság szignifikanciaszintje némileg meghaladja a konvencionális 10 százalékot, akkor azt mondhatjuk, hogy közel 2 százalékos jövedelemcsökkenést okoz a Budapesttől való vasúti távolság az átlagos településnek.

4. Tanulságok

A tanulságokkal kapcsolatban előzetesen le kell szögezni, hogy minden számszerű következtetésünk, hasonlóan a többi, témában megnyilatkozó tanulmányhoz, igen korlátozott kereteken belül használható fel gazdaságpolitikai vagy stratégiai célokra. Oksági összefüggések megállapítására megfelelő idősorok hiányában ma Magyarországon nem vállalkozhatnak a téma kutatói, így én sem. Az ebben a tanulmányban követett stratégia sem az okság bizonyítására vagy elvetésére vállalkozik, hanem annak bemutatására, hogy a jelenleg létező hazai tanulmányok elérhetőségi vizsgálatai fontos változót felejtenek el, mikor a közlekedési infrastruktúra elérhetőségi-mutatókat kizárólag autópálya-mutatóként értelmezik, illetve e mutatók jelentőségét csak az autópálya-fejlesztés melletti érvként hozzák fel.

Megállapításainkat tehát ennek tükrében kell értékelní. Ha jók a közlekedési infrastruktúrával kapcsolatos számítások, akkor is látható, hogy a települések hátránya

teljes egészében nem szüntethető meg. Aligha jó az a megközelítés, hogy amennyiben a település és Budapest közötti távolságot *teljes egészében* autópályán lehetne megtenni, akkor mennyivel javulna a foglalkoztatás és a jövedelem. A geográfiai távolság olyan adottság, amelyet az útminőség javítása csak *részben* szüntethet meg, ezért nem tetszés szerinti a javulás lehetősége.

A vizsgálat eredményeinek összefoglalása

Változó	Teljes minta				Változó	Részminta			
	Foglalkoztatás		Jövedelem			Foglalkoztatás		Jövedelem	
	együtthatók					együtthatók			
	felülről építkező	alulról építkező	felülről építkező	alulról építkező		felülről építkező	alulról építkező	felülről építkező	alulról építkező
Kozut02	-0,177127	-0,186063	-	-0,360308	Vasutav	-0,143047	-0,204282	-	
Sq_kozut	0,000674	0,0006443	-		Sq_vasutav	0,0004831	0,0006312	-	-0,0006472
A távolság szempontjából kifejtett hatás									
	százalékpont		ezer forint			százalékpont		ezer forint	
	FoglB	FoglB	szja01	szja01		FoglB	FoglB	szja01	szja01
Átlag	-10,5	-11,9	-	-40,6	Átlag	-9,6	-15,0	-	-14,5
Medián	-11,5	-13,0	-	-42,8	Medián	-10,6	-16,3	-	-13,4
Minimum	-6,2	-6,6	-	-14,9	Minimum	-6,0	-8,8	-	-1,7
Maximum	-7,6	-10,7	-	-75,4	Maximum	-4,5	-10,4	-	-43,8
Százalékos hatás									
Átlag	-20,2	-22,9	-	-5,2	Átlag	-18,3	-28,7	-	-1,8
Medián	-20,2	-22,8	-	-5,9	Medián	-21,0	-32,4	-	-1,7
Minimum	-10,3	-11,0	-	-1,5	Minimum	-11,2	-16,3	-	-0,2
Maximum	-32,9	-46,7	-	-12,0	Maximum	-8,3	-19,0	-	-6,0

Megjegyzés. A becsléseket OLS-módszerrel számítottuk. A kozut02 a Budapesttől mért közúti távolság, az útszakaszok minőségével súlyozva; az Sq kozut02 ezen mutató négyzete; a vasutav a Budapest vasúti eléréséhez szükséges minimális idő; az Sq vasutav ezen mutató négyzete; a FoglB az adott településen élők közül a foglalkoztatottak/a település népessége-településen élő munkanélküliek-inaktívok; az szja01 a jövedelmi mutató, a településen bevallott összes adóalap/a bevallók száma.

Úgy tűnik, hogy az autópálya továbbépítése egy 80 kilométeres átmérőjű körön belül a foglalkoztatást mintegy 4 százalékos mértékben javíthatja, és hasonló mér-

tékben a jövedelmeket is. De nem találunk bizonyítékot a 80 kilométeres átmérőjű körön belüli közelebb hozás pozitív hatására. Ez az eredmény ellentmondásban van azzal a megállapítással, hogy az autópályák csak közvetlen környezetükben tudnak önmagukban jövedelemnövelő hatást kiváltani, néhány tíz kilométerrel a csomópontoktól ez a hatás elvész.²⁵ Nincs összhangban továbbá azokkal az elemzésekkel, amelyek az autópályák 30 kilométeres belső és 30-50 kilométeres külső sávját összevetve találják úgy, hogy a pozitív hatások csak a belső sávra koncentrálnak (*Németh* [2004] 163-166. old). Az összhang hiányát esetleg az magyarázza, hogy a mi elérhetőségi mutatónk nem fizikai távolságot, hanem utazási időt vet össze. Más tanulmányok egyes eredményei sem látszanak igazolni egyébként teljes mértékben, az autópályától való távolság döntő fontosságát. Így *Németh* [2004] tanulmányának 7. ábrájából, amely az autópályák belső sávját hasonlítja össze a külsővel az 1990-es években, nem látszik, hogy a belső sávban valóban kedvezőbben változott volna a munkanélküliség. Az M5-ös autópálya esetében mind a belső, mint a külső sávban a relatív munkanélküliségi mutató ugyanolyan mértékben látszik csökkenni. Az M3-as autópálya belső sávjában a munkanélküliség kevésbé esett, mint a külsőben 1990 és 2002 között.²⁶

A vasúti elérhetőség biztosítása annyi haszonnal járna a foglalkoztatásban és a jövedelmi helyzetben, mint az autópálya 80 kilométeres átmérőjű körön belülré hozása.

Az általunk használt mutatók javulása úgy járulhat hozzá – feltevésünk szerint – a foglalkoztatás és a jövedelmek növekedéséhez, hogy ezáltal csökkenne az utazási idő. Ez nemcsak a közúton, hanem a vasúton való gyorsabb elérés melletti érvként is értelmezhető, ami nyilván az utóbbi fenntartására és karbantartására fordított nagyobb erővel érhető el.

Eredményeim egyfelől tulajdonképpen több ponton harmonizálnak az autópálya-építések *mellett* empirikus számítások alapján érveket találókéval. Az autópályák pozitív hatását nem vonják kétségbe. Viszont alapvetően más jellegűek, mivel felhívják a figyelmet arra, hogy az elérhetőség javítását nem lehet kizárólag az autópályára szűkíteni. A következtetések legalább annyi érvet szolgáltatnak a vasút fejlesztésére, mint amennyit a többi tanulmány az autópályák bővítésére. (Ugyanakkor mindannyiunkat terhel az oksági elemzés hiánya.) Tanulmányom nem a „vasút iránti nosztalgián”, vagy valamiféle „közútszkepticizmuson” alapszik. (*Németh* [2004] 149. old.). Ugyanis *mindkét* közlekedési móddal szemben „méltányosan” jár el, ha ennek a kifejezésnek van értelme ebben az összefüggésben. Ez a „méltányosság” viszont szerintem hiányzik még a hozzám hasonló módszert al-

²⁵ Németh ezzel magyarázza, hogy a több tíz kilométeres sávokat felölelő kistérségek szintjén miért tűnik el a települési szinten még érezhető pozitív autópálya-hatás (*Németh* [2004] 153. old.).

²⁶ Mindazonáltal a relatív jövedelmi helyzet valóban kedvezőbben változott az M3-as és az M5-ös autópálya belső sávjában, mint a külsőben *Németh* [2004] 6. ábrája alapján.

kalmazó, itt vizsgált tanulmányokból is. Az autópályák foglalkoztatás- és jövedelemnövelő hatását nem tagadjuk, és nem is kell tagadni. Az alapvető probléma az, hogy az autópálya meglétével annak az *elérhetőségre* gyakorolt pozitív hatását mérik, amit viszont a vasút és a hagyományos közút is képes biztosítani. Az elérhetőségnek viszonylag jó proxyja az autópálya, de jó a vasúti kapcsolat léte is, sőt még jobb a közúti és a vasúti távolság.

A települések hátrányának csökkenése a település elhelyezkedésétől, a fejlesztett közúti és vasúti szakaszok pályájának megválasztásától függően széles sávban ingadozhat; összességében azonban a különbség 10 százaléknak a megszüntetése (a *kozut02* változó értékének 10 százalékos csökkentése) már igen jó eredménynek lenne mondható.²⁷ Ez alapján úgy tűnik, hogy a várt foglalkoztatási haszon csak igen kedvező esetben haladhatja meg az 1 százalékpontot, a jövedelmi haszon is alacsony, 1 százalékon belüli.

Eredményeim az alsó határa alatt vannak annak a jövedelemnövekménynek, amit *Kullman* [1999] kimutattott. Ő 5 000-15 000 forintot egy főre jutó jövedelemnövekményt számított 1997-es árakon, viszont én egy optimista 10 százalékos elérhetőségjavulás esetén is csak 4 000 forintot, ráadásul 2001-es árakon. *Kullman* számításainak alapulvételével az autópálya építési költségeinek 15 éves megtérülése adódott az állam számára, a mi eredményünk szerint ez az idő mintegy háromszor hosszabb.

A közúti és a vasúti távolság négyzete is szignifikáns, ami miatt a legtávolabbi mintabeli településen kisebb az *abszolút* foglalkoztatási *elmaradás* az átlagosnál (a felülről építkező modellben a közút esetében 1-3, a vasút esetében 5 százalékponttal). Ez az eredmény megerősíteni látszik azt a mások által képviselt álláspontot (*Tóth* [2005] 10. old.), hogy az autópálya a legszegényebb/legtávolabbi településeken a foglalkoztatást illetően átlag alatti mértékben segít. Viszont ellentmond azoknak, akik szerint az autópálya fejlesztési hatása annál erősebb, minél távolabb van egy kistérség Budapeستől és a nyugati határtól (*Bartha-Klauber* [2000]).

Jacoby [2000] 10 százalékos körüli jövedelemnövekedést számszerűsített a közúti elérhetőség megerősítése esetén egy fejlődő ország, Nepál esetében. Az általunk kapott eredményt ezzel összevetve a közúti fejlesztéstől várt 1 százalékon belül maradó haszon igen szerénynek tűnik. A viszonylag szerény mértékű javulás arra vezethető vissza, hogy mi már nem vagyunk az a szegény ország, ahol a fizikai infrastruktúra megerősítésével többszörösére emelkedhet a szállítási sebesség, illetve törtrészére csökken a szállítási költség. Nálunk a közúti és a vasúti infrastruktúra tulajdonképpen minden települést bekapcsol az országos hálózatba. A közúti és vasúti infrastruktúra mennyisége tekintetében (minőségére ez nem igaz) Magyarország nem marad el

²⁷ Az átlagos település 76 kilométer autópályán, 40 kilométer elsőrendű főúton, 24 kilométer másodrendű főúton és 8 kilométer egyéb úton érhető el Budapeستől. Ez alapján a *kozut02* mutató értéke 112. Ha az *elsőrendű főút autópályává alakul*, akkor a mutató értéke 106-ra, 6 százalékponttal csökken. A *kozut02* mutató alapján maximális távolságra levő település (Cigánd) esetében a relatív csökkenés ennél valamivel kisebb.

az európai országok átlagától. Autópályák – különösen igaz ez a további autópályákra – építésével csak a már egyébként is viszonylag magas sebességű, gépesített közlekedést lehet még valamelyest gyorsítani, és csak azokon a szakaszokon, ahol még nem autópálya, illetve az elhanyagolás miatt leromlott/lelassult vasút biztosítja a szóba forgó település elérhetőségét.

Ha a közlekedési elérhetőség javulásától várható foglalkoztatási hasznokat összevetjük a helyi tömegközlekedés fejlesztésétől *Kertesi-Köllő* [1998] által várt hasznokkal, akkor jelentős különbséget mutathatunk ki az utóbbi javára. Ők, mint arról volt szó, 5-6 százalékpontos különbséget mutattak ki a jó és a rossz tömegközlekedési adottságú települések között. Ezzel szemben az általunk számolt, az országos léptékű elérhetőség javításából (vasúti és közúti forgalom gyorsítása) fakadó foglalkoztatási többlet 1 százalékpont alatti.

Az egyéb, nem közlekedési infrastruktúrára vonatkozó eredmények jórészt alátámasztják a tanulmány elején számba vett, más szerzőktől származó empirikus eredményeket. A kistérségi jellemzők jelentős mértékben befolyásolják a települési szintűeket. Az ingázás kedvező hatását mi is ki tudtuk mutatni. A vállalkozássűrűség a mi számításainkban is pozitívan hat a foglalkoztatásra, az alacsony iskolai végzettség pedig negatívan (amint a jövedelemre is). Kisebb mértékben meglepő a részvénytársaságok és a termelészövetkezetek sűrűségének indifferenciája. Úgy tűnik, hogy a távolsági buszkapcsolat léte sem perdöntő a többi közlekedési jellemző mellett. A város-falu különbség hatását pedig a többi tényező általában lefedi, így a település jellegének önmagában, elvonatkoztatva a modellbe épített tényezőtől, nem szignifikáns a hatása (bár a vasúti távolság-változó alkalmazásakor, a felülről építkező, a foglalkoztatást magyarázó modellben szignifikáns volt a településtípus-változó, a várt előjellel).

Furcsa, de talán nem megmagyarázhatatlan a vállalkozássűrűség eltérő hatása a foglalkoztatásra és a jövedelmekre. Figyelmeztetni kell arra, hogy a jövedelmekre gyakorolt negatív hatás – az átlagos település esetén – a modelltől függően igen széles sávban (0,5-14%) mozog, ezért amennyiben az alsó becslés áll közelebb az igazsághoz, akár kicsinek is minősíthető. A jövedelmi mutatót az szja-bevallások alapján készítettük, viszont a vállalkozók nem jeleskednek ennek az adónak a fizetésében. Így a szélesebb adóalapon közelített jövedelmi haszon talán más eredményt adna.

Ami nyugtalanító, hogy a települések ágazati foglalkoztatottsági szerkezete és a felsőfokú végzettségük aránya – több modellben – a várttal ellentétben viselkedik, ami a foglalkoztatást illeti. Ez azért is meglepő, mivel a nemzetközi irodalomban és a bemutatott magyar tanulmányokban is, a humán tőke igen jelentős mértékben pozitívan járul hozzá a növekedéshez és a termelékenység javulásához.²⁸ Az egyetemi végzettség aránya talán azért bizonyulhat inszignifikánsnak vagy érintheti esetleg negatívan a foglal-

²⁸ *Németh* ([2004] 153. old.) hasonló következtetésre jut: az ő iskolázottsági mutatójának együttthatója messze meghaladja az autópálya-hálózatához képest való elhelyezkedés és a nyugati határszéltől való távolság mutatóját.

koztatottságot, mert minél kisebb a foglalkoztatási ráta, annál inkább csak az egyetemet végzettek azok, akik munkához jutnak. A felsőfokú képzés ugyanakkor a jövedelmek esetében az elmélet alapján a vártnak megfelelően – akárcsak az ágazati szerkezet – nagyon kedvező hatást gyakorol. Ez is megerősíti a több kutató által említett, humántőke fejlesztésre alapozó gazdaságpolitika létjogosultságát.

Eredményeink ellen fel lehet hozni, hogy azok csak két autópályára vonatkoznak, és mint egyesek megállapítják (Németh [2004] 169. old.) nem lehet egységesen kezelni az autópályákat a tőlük várt térszerkezet-alakító hatás szempontjából. Azzal valószínűleg magam is egyetértek, hogy az autópályák nyomvonalválasztása aligha független attól az előnytől, amit tőlük várni lehet. Csakhogy akik ezt megállapítják, olykor maguk tesznek olyan általános megállapításokat, amely az autópályák általában vett hasznosságáról szólnak. Véleményem szerint minden konkrét regionális szállítási infrastruktúrafejlesztés esetében, amennyiben elfogadjuk, hogy az elérhetőség javításával kívánunk hozzájárulni a jövedelmek és a foglalkoztatás növeléséhez, az infrastruktúra minden komponensével foglalkozni kell.

Irodalom

- BARTHA A. – KLAUBER M. [2000]: Az M5 autópálya gazdasági hatásvizsgálata. *Ipari Szemle*. 20. évf. 4. sz. 20–22. old.
- BÍRÓ P. – MOLNÁR L. [2004]: A kistérségek fejlettségi szintjének és infrastruktúrájának összefüggései, *Közgazdasági Szemle*. 51. évf. 11. sz. 1048–1064 old.
- EUROPEAN COMMISSION [2003]: Public Finances in EMU. In: *European Economy*. Brüsszel.
- FLEISCHER T. ET AL. [2002]: *A Széchenyi terv autópálya fejlesztési programjának stratégiai környezeti hatásvizsgálata*. Környezettudományi Intézet. Budapest.
- GKI [2003a]: *Az autópálya makrogazdasági hatásainak modellezése, EU területfejlesztési politikai összefüggések*. Budapest.
- GKI [2003b]: *A kistérségek fejlettségi szintjének és infrastruktúrájának összefüggései*. Budapest.
- KERTESI G. [2000a]: A cigány foglalkoztatás leépülése és szerkezeti átalakulása. *Közgazdasági Szemle*. 47. évf. 5. sz. 406–443. old.
- KERTESI G. [2000b]: Ingázás a falusi Magyarországon. *Közgazdasági Szemle*. 47. évf. 10. sz. 775–798. old.
- KERTESI G. – KÖLLŐ J. [1998]: Regionális munkanélküliség és a bérek az átmenet éveiben. *Közgazdasági Szemle*. 45. évf. 7–8. sz. 621–652. old.
- KISS K. – LUKÁCS A. (szerk.) [2003]: *Uniós csatlakozás– közlekedés– környezet*. Levegő Munkacsoport. Budapest.
- KÖLLŐ J. [1997]: A napi ingázás feltételei és a helyi munkanélküliség Magyarországon. *Esély*. 8. évf. 2. sz. 33–61. old.
- KULLMAN, Á. [1999]: Kísérlet a Füzesabony-Polgár autópályaszakasz területfejlesztő hatásának számszerűsítésére. *Falu-Város-Régió*. 7. sz. 18–20. old.

- KSH [2004]: 1990. és 2001. évi népszámlálás. Területi adatok, Bács-Kiskun, Borsod-Abaúj-Zemplén, Csongrád, Heves megye. www.nepszamlalas.hu
- JACOBY, H. [2000]: Access to markets and the benefits of rural roads. *The Economic Journal*. 110. évf. 465. sz. 713–737. old.
- LEVEGŐ MUNKACSOPORT [2003]: *Az államháztartás ökoszociális reformjának szükségessége és lehetőségei. Ajánlások a 2004. évi állami költségvetéshez*. Levegő Munkacsoport. Budapest.
- NÉMETH N. [2004]: Az autópálya-hálózat térszerkezet alakító hatásai – Magyarország esete. In: *A hely és a fej. Munkapiac és regionalitás Magyarországon*. Közlekedéstudományi Intézet. Budapest.
- OHNSORGE-SZABÓ L. – KAJNER P. – UNGVÁRI G. [2005]: *Fenntartható EU-felé*. L'Harmattan. Budapest.
- OHNSORGE-SZABÓ L. [2005]: Infrastruktúra-fejlesztés gazdasági hatása. *Statisztikai Szemle*. 83. évf. 8. sz. 737–759. old.
- SZALKAI G. [2001]: Elérhetőségi vizsgálatok Magyarországon. *Falu, város, régió*. 10. sz. 5–13. old.
- TÓTH G. [2005]: Potenciálmodell alkalmazásának lehetőségei az autópálya-nyomvonalak területfejlesztési szempontú vizsgálatában. *Gazdaság és Társadalom*. 17. évf. 3. sz. 3–17. old.
- VÖRÖS A. – POLÁNYINÉ CSÁNYI Á. – CZEGLÉDI L. [2003]: Az M0-s autótérület- és gazdaságfejlesztő hatásának figyelemmel kísérése kérdőíves adatfelvétel alapján. *Városi Közlekedés*. 43. évf. 3.sz. 136–140. old.

Summary

The purpose of this study is to analyse the regional developmental effect of motorways, on the assumption that this effect depends on the time accessibility of the region concerned. The method of other empirical studies arguing for enlargement of the motorway network in Hungary is also criticized. It is shown that though these studies explored many relevant facts and relationships, even the methodologically most sophisticated ones can be used only in a limited way for development- and regional policy decisions. They are not able to establish causal relationship between motorway and wealth. This study cannot undertake for establishment of causal relations as well, but wishes to verify that Hungarian researchers have narrowed their interests on the motorways were without sufficient reasons. Has one investigated the transportation infrastructure in the broad sense, including public roads and railroad system, can one found the one-sided priority given for the motorways is not well-founded.

Az anyagáramlás-elemzés (statisztikai) módszertani kérdései I.

Szabó Elemér

a Környezetvédelmi
és Vízügyi Minisztérium
főtanácsosa

E-mail: szabo@mail.kvvm.hu

Pomázi István

a Környezetvédelmi
és Vízügyi Minisztérium
szakmai főtanácsadója

E-mail: pomazi@mail.kvvm.hu

A nemzetgazdasági anyagáram-elszámolás (Material Flow Accounts/Accounting – MFA) fő célja összegzett háttérinformáció biztosítása a társadalmi-gazdasági rendszer fizikai szerkezetének összetételéről és változásairól. Az anyagáram-elszámolás hasznos eszköz a gazdaság és a környezet kölcsönhatásának elemzéséhez, továbbá a környezeti és integrált környezeti, a társadalmi és gazdasági mutatók származtatásához. Ezek a mutatók lehetővé teszik a bruttó hazai termékhez (GDP) hasonló összegzett gazdasági mutatókkal való összehasonlítást és így segítenek a politika figyelmét a tisztán pénzügyi elemzéstől a biológiai-fizikai szempontok beépítése felé terelni.

Az összefüggő és átfogó adatszervezésnek köszönhetően az anyagáram-elszámolás közvetlenül kapcsolódhat a létező gazdasági elszámolási rendszerekhez, például a nemzeti számlarendszerhez, és része lehet a kiterjesztett környezeti és gazdasági számláknak, ilyen például az ENSZ környezeti-gazdasági elszámolási rendszere.

A megszokott anyagáram-elszámolási módszernek két fő hiányossága van: a különböző nagyságú anyagáramlások szintetikus (összegzett) mutatók előállítására történő összegzése, valamint az anyagáramlás-mutatók és a környezeti hatások közötti bonyolult és áttételes kapcsolat.

TÁRGYSZÓ:
Környezetstatisztika.

Az Európai Unió 2002-ben elfogadott, tíz évre szóló Hatodik Környezetvédelmi Akcióprogramja egyik fő célként fogalmazza meg a fenntartható erőforrás-használatot, különös tekintettel a nem megújuló erőforrásokra. A fő cél az erőforrás-használat szétválasztása a gazdasági növekedéstől a jelentős mértékben javuló erőforrás-hatékonyság a gazdaság dematerializációja és a hulladékmegelőzés révén. Az Akcióprogram a később kidolgozandó tematikus területek egyikeként a természeti erőforrások fenntartható használatával foglalkozó stratégiát jelölte meg. A tematikus stratégiáról 2003-ban kiadott Bizottsági Közlemény fontos környezetpolitikai eszközként nevezte meg az anyagáram-elszámolásokat és -elemzéseket. Az ENSZ 2002-ben, a dél-afrikai Johannesburgban megtartott Fenntartható Fejlődés Világtalálkozón elfogadott Végrehajtási Tervben arra hívta fel a tagállamok figyelmét, hogy ösztönözzék egy tízéves fenntartható termelési és fogyasztási keretprogram kidolgozását, amelynek célja a hatékonyság és a fenntarthatóság javítása az erőforrások felhasználásában.

Az anyagfelhasználás intenzitása globálisan növekszik a termelés ésszerűsítésén és a szolgáltatásokon, valamint az infokommunikációs technológián alapuló globális gazdaság részleges dematerializációja ellenére. Ezért szükséges a gazdasági folyamatok hatékonyságának mérésére szolgáló módszerek keresése. Az anyagáramlás-elemzés az egyik olyan módszer, amely kellőképpen fel tudja tárni a társadalmi metabolizmusként ismert jelenség folyamatait és összefüggéseit.

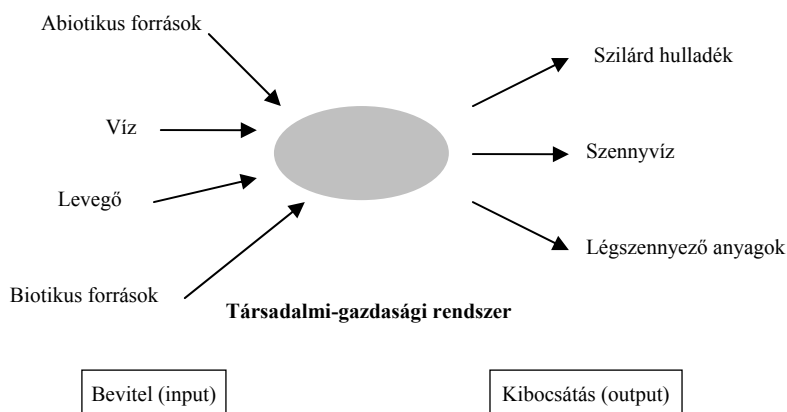
A természetben zajló folyamatok sebességéhez képest gyors változások következménye lehet rövid és hosszú távon egyaránt negatív hatásainak megjelenése és felerősödése a gazdaságkörnyezeti rendszerben, ami veszélyeztetni fogja az ember fejlődését és túlélését. Ezért a gazdasági rendszeren keresztül folyó természetes eredetű anyagáramlásokat fel lehet használni a környezetterhelés és -fenntarthatóság mérésére alkalmas mutatóként. A pénzügyi mutatókkal (például GDP) összevetve, az anyagáramlás mint fizikai mutató alkalmasabb a különböző régiók és különféle időszakok fenntarthatóságának mérésére.

1. Az anyagáram-elszámolás módszertani alapjai

Az anyagáram-elszámolás (Material Flow Accounts/Accounting – MFA) megfelelő módszer lehet egy gazdasági rendszer nem fenntarthatóságának mérésére olyan szempontok figyelembevételével, amelyeket a gazdaság pusztán pénzügyi szemléletű vizsgálata teljes mértékben figyelmen kívül hagy. Az MFA nemzetközileg elfogadott kulcs-

fontosságú eszköz a társadalom biofizikai anyagcseréjének elemzésében, és összevont mutatókat biztosít az emberi tevékenységek környezetterhelésének méréséhez.

I. ábra. Az anyagáram-elemzés vázlatos modellje



Forrás: Hinterberger F. et al. [2003].

A nemzetgazdasági MFA-megközelítés alap elképzelése a gazdaság és a környezet közötti kölcsönhatás olyan egyszerű modellje, amelyben a gazdaság a környezetbe beágyazott alrendszer, és – hasonlóan az élőlényekhez – az anyagok és az energia folytonos átáramlásától függ. Az alapanyagok, a víz és a levegő mint bemenő (input) anyagok a természetből kerülnek ki és termékekké alakulnak át, végül mint kimenő (output) anyagok (hulladék és szennyezőanyag-kibocsátás) visszakerülnek a természetbe. A természetes anyagcsere-folyamatok hasonlóságának kiemelésére az „ipari” (Ayres [1989]) vagy „társadalmi” (Fischer-Kowalski [1998]) anyagcsere fogalmát vezették be.

A környezet és a gazdaság kölcsönhatásainak leírására alkalmas anyagmérlegeket a hatékony használat érdekében anyagcsoportonként és gazdasági ágazatonként kell szétbontani. Az anyagcsoportok megfelelő szintű szétbontását tartalmazzák az elsődleges anyagbeviteli statisztikák, míg a gazdasági ágazatok szétbontására a gazdaságstatisztikák, a hulladékstatisztikák és az energia- légszennyezés-kibocsátási statisztikák alkalmasak.

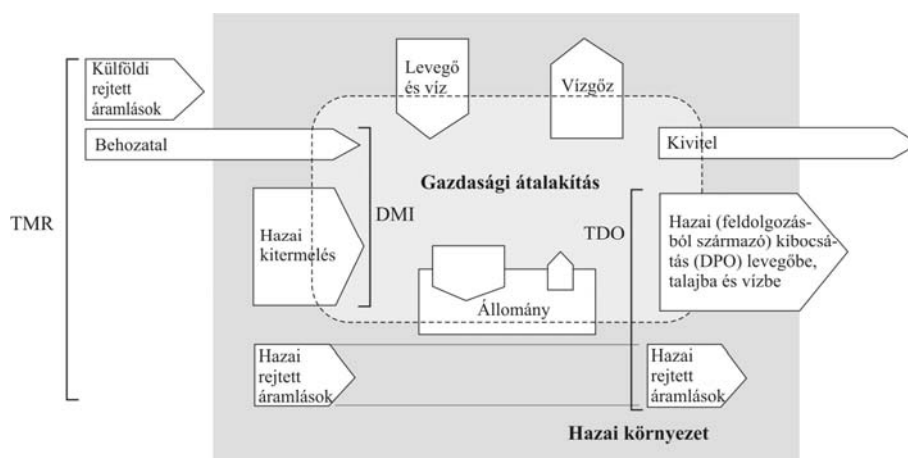
Az összanyagigényletet (Total Material Requirement – TMR) mint keretrendszert gyakran éri bírálat, amiért az anyagbevitelt tonnában méri, mivel ez nem mond sokat a természeti erőforrások használatából fakadó környezeti hatásokról. Ugyanilyen kritika illeti az anyagmérlegeket is. Nem világos, hogy mi lenne az anyagmérlegek alkalmazása és politikai lényege, és vajon javítanák-e a környezeti-gazdasági döntéshozatal adat- és információalapját. Valamennyi bevitel és kibocsátás részletes

és ésszerű anyagmérlegének összeállítása és elemzése nagyon költséges, ezért ritkán kap kiemelt szerepet a statisztikában és a statisztikai rendszerek fejlesztésében.

Mindenesetre a TMR és az anyagmérlegek épp úgy érdekes és fontos információkat szolgáltatnak politikai, kutatási és modellezési célokra, mint a közvetlen anyagbevitelről és a rejtett áramlásokról. A TMR és az anyagmérlegek ágazonként összekapcsolhatók mind a gazdasági statisztikákkal, mind a környezeti számlát magába foglaló Nemzeti Elszámolási Mátrixszal (National Accounting Matrix including Environmental Accounts – NAMEA). A TMR manapság már kiegészül a hulladékokkal és kibocsátásokkal, valamint a bevételhez és a kibocsátáshoz kapcsolódó gazdasági kérdésekkel, noha nem közös egységekben kifejezett adatok kerülnek egymás mellé.

Az anyagok beviteli és kimenő oldalának kiegyensúlyozásakor azok közös egységekben történő kifejezésével merül fel a legtöbb probléma. Sok esetben akkor érhető el a legjobb információ, ha az anyagok bemenő és kimenő oldalán létező statisztikákat ágazonként és anyagcsoportonként – de nem szükségszerűen tonnában kifejezve – mutatják be.

2. ábra. Az anyagáram-elszámolás és -elemzés alapmodellje



Megjegyzés. A rövidítések magyarázatát (kifejtését) lásd a 281. oldalon.

Forrás: Matthews, E. et al. [2000].

A környezet és gazdaság keretrendszerét célzó statisztikák összekapcsolhatóságának fejlesztése kiemelt fontosságú. A mérlegek nem szükségszerűen terjednek ki az összes lehetséges anyagra és ágazatra. Egyik lehetőség a kiválasztott és a környezetre veszélyt jelentő nagy mennyiségű anyagokra vonatkozó mérlegek meghatározása. A kulcsfontosságú anyagok azok lehetnek, amelyeknek nyilvánvaló környezeti, továb-

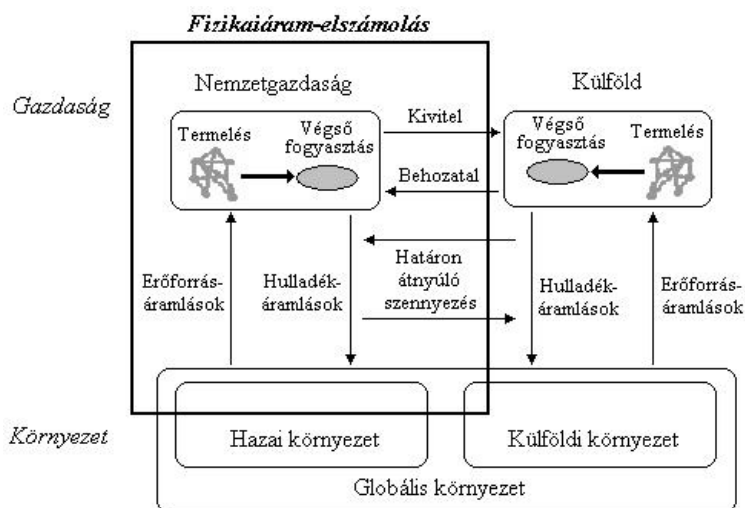
bá a földhasználattal – különösen a kitermeléstől és az elsődleges bevittől a végső fogyasztásig és környezetbe való visszajuttatásig terjedő környezeti következményekkel – összefüggő hatásaik vannak.

A termodinamika első főtétele (energiamegmaradás) alapján egy rendszerben az összes bevitelnek meg kell egyeznie az összes kibocsátással és a nettó anyaggyarapodás összegével. Ez az anyagmérlegelv igaz a gazdaságra mint egészre és bármely alrendszerére is.

Egy nemzetgazdasági szintű anyagáram-számla összeállítása céljából pontosan kell meghatározni a gazdasági és környezeti rendszer határait, hiszen csak a határokat átlépő erőforrások kerülnek figyelembevételre. Az ENSZ integrált környezeti és gazdasági számlarendszerében (System of Integrated Environmental and Economic Accounts – SEEA) leírtak szerint a gazdasági rendszert a hagyományos nemzeti számlarendszer (System of National Accounts – SNA) által figyelembe vett folyamatok határozzák meg. Így az SNA három gazdasági tevékenységtípusához (termelés, fogyasztás és készletváltozás) kapcsolódó valamennyi folyamat a gazdasági rendszer részét képezi, a környezeti rendszer pedig magában foglalja a piacon forgalmazott termékeken kívül az összes erőforrást.

Ezért a nemzetgazdasági MFA esetében az erőforrás-áramlásokra két fő határ nevezhető meg. Az egyik a gazdaság és a hazai környezet közötti határ, amelyből az erőforrások (anyagok, víz, levegő) származnak, a másik a többi gazdasággal folytatott kereskedelem mint áramlás határa.

3. ábra. Fizikai áramlások és a fizikaiáram-elszámolás köre



Forrás: Eurostat [2001].

2. Az anyagáramlások osztályozása és elemzési típusai

Egy országos szintű, átfogó anyagszámítás-rendszer felvázolása előtt tisztázni kell az egyes anyagáramlások közötti különbséget. Az Eurostat [2001] módszertani útmutatója alapján a következő anyagáramlás-típusokat lehet megkülönböztetni.

a) *Közvetlen és közvetett anyagáramlások.* A közvetlen áramlások a termék valódi tömegére vonatkoznak, így nem veszik figyelembe a termelési lánc életciklus-dimenzióját. A közvetlen áramlások azonban megmutatják az összes anyagot, amely a feldolgozás (felmenőerőforrás-igény) során szükségessé válik, és egyaránt magukban foglalják a felhasznált és a fel nem használt anyagokat.

b) *Felhasznált és fel nem használt anyagáramlások.* A felhasznált anyagok csoportját azon kitermelt erőforrások mennyisége határozza meg, amelyek további feldolgozásra vagy közvetlen fogyasztás céljára a gazdasági rendszerbe lépnek. A gazdasági rendszerben valamennyi felhasznált anyag átalakul vagy átalakítják. A fel nem használt kitermelt anyag az az anyagszámítás, amely sohasem kerül be a gazdasági rendszerbe, így a fizikai piac externáliájaként írható le (*Hinterberger-Luks-Stewen* [1996]). Ez az osztály magában foglalja a bányászati tevékenységből származó takarórétegeket és elválasztó meddőrétegeket, a biomassza-kitermelésből származó, nem szándékolt halászati fogásokat és fakitermelési veszteségeket, valamint az építőipari tevékenységekből származó földkiemelést és kotort anyagot.

A fel nem használt és közvetett anyagáramlások osztályára a nemzetközi irodalomban az *ökológiai hátizsák* (*Schmidt-Bleek* [1994]) vagy a *rejtett áramlások* (*Adriansee et al.* [1997]) kifejezést is gyakran használják.

c) *Hazai és külföldi anyagáramlások.* Ez az osztály az áramlások eredetére és/vagy végállomására vonatkozik.

Az említett három dimenzió alapján a nemzetgazdasági MFA számára fontos anyagbevitel öt osztályát az 1. táblázat mutatja be.

1. táblázat

A nemzetgazdasági MFA anyagbevitelének osztályai

Anyag	Gazdasági kezelés	Eredet	Használt fogalom
Közvetlen	Felhasznált	Hazai	Hazai kitermelés (felhasznált)
<i>(Nem alkalmazható)</i>	Fel nem használt	Hazai	Fel nem használt hazai kitermelés
Közvetlen	Felhasznált	Külföldi	Behozatal
Közvetett	Felhasznált	Külföldi	A behozatalhoz kapcsolódó közvetett bevitel
Közvetett	Fel nem használt	Külföldi	áramlásai

Forrás: Eurostat [2001].

Az egyszerű anyagáram-számla a szilárd anyagok áramlására összpontosítja figyelmét. Ez a csoport további három alcsoportra bontható:

- ásványok (fémércsek és nem fém ásványi anyagok, például kő, agyag stb.),
- fosszilis energiahordozók (kőszén, kőolaj és földgáz)
- biomassza (mezőgazdaságból, erdőgazdálkodásból és halászatból).

d) Anyagállomány. A fizikai elszámolás szempontjából a nagy mértékű fizikai állomány-gyarapodás a modern ipari társadalmak egyik fő jellegzetessége. Az MFA-keretben az állomány főleg az emberi eredetű állományt jelenti: egyfelől az infrastruktúra és épületek, másfelől a tartós fogyasztási cikkek (autók, háztartási berendezések) és beruházási javak (gépgyártás). Az erdők és a mezőgazdasági haszonnövények a környezeti rendszer részét képezik, ezért nem kerülnek a fizikai állományba, de a fakitermelés és terménybetakarítás mint bevétel megjelenik a gazdasági rendszerben. Az ellenőrzött lerakókon elhelyezett hulladékok a gazdaság szempontjából inkább környezeti kibocsátásként, semmint fizikai állományként jelennek meg.

3. Az ipari metabolizmus elemzésének szintjei és típusai

Az ipari metabolizmus szerkezetének (Ayres [1989]; Náray-Szabó [1999]), mennyiségének és minőségének megértése függ az anyagáramlás-elemzések mikéntjétől, azaz attól, hogy a folyamat melyik szakaszára terjed ki az erőforrás-kitermeléstől a végleges hulladék-elhelyezésig. Az elemzéseknek különböző szintjeit lehet megkülönböztetni:

- termékek és szolgáltatások értékelése az életciklus alapján. Ebben az esetben az életciklus-elemzés biztosít átfogó képet, például a kumulatív energiaigények vagy az egységnyi szolgáltatásra jutó anyagbevitel számszerűsítésére;
- a vállalatok esetében az inputok és outputok fizikai mérlegét egyre inkább alkalmazzák a környezeti teljesítmény jelentés részeként, és ez lényeges információkat szolgáltat egy adott vállalat környezeti irányításáról;
- az adott ágazatok szintjén is jól alkalmazható az anyagáramlás-elemzés mindkét vertikális irányban (alulról felfelé vagy fordítva);
- közösségek (települések), régiók és nemzetgazdaságok anyagáramlásainak vizsgálata egyre fontosabbá válik a különböző politikai döntések megalapozásában.

Az anyagáramlás-elemzéseknek alapvetően két fő típusa különböztethető meg: az „I”-típusú vizsgálatok inkább technikai-technológiai szempontokat vesznek figyelembe, míg a „II”-típusú elemzések a gazdasági-társadalmi vonatkozásokra irányulnak. Ezek a fő típusok nem esnek szigorúan egybe a korábban bemutatott két stratégiai iránnyal. A detoxifikációs koncepció (mérgező anyagok kivonása) ugyanakkor erősebben van jelen az „Ia”-típusnál, míg a dematerializációt világosabban képviseli a „IIc”-típus.

2. táblázat

Az anyagáramlásra vonatkozó elemzések főbb típusai

Vizsgálati témakörök	Az elemzés típusa					
	I			II		
	a	b	c	a	b	c
	Egyes környezeti problémákhoz kapcsolódó hatások*			Az átáramláshoz kapcsolódó környezeti problémák		
	Kémiai elemek	Anyagok	Termékek	Vállalatok	Ágazatok	Régiók
	például Cd, Cl, Pb, Zn, Hg, N, P, C, CO ₂ , CFC	például faipari termékek, energia-hordozók, földkiemelés, biomassza, műanyagok	például pelenkák, akkumulátorok/elemek, autók	például egyedi üzemek, közepes és nagyvállalatok	például termelő ágazatok, vegyipar, építőipar	például teljes körű vagy főbb átáramlások, tömegáramlás-mérleg, összanyagigényesség
	Egyes vállalatokon, ágazatokon és régiókon belül			Kémiai elemekhez, anyagokhoz és termékekhez kapcsolódóan		

* Egységnyi áramlásra vetítve.

Forrás: Brinzeu [2000].

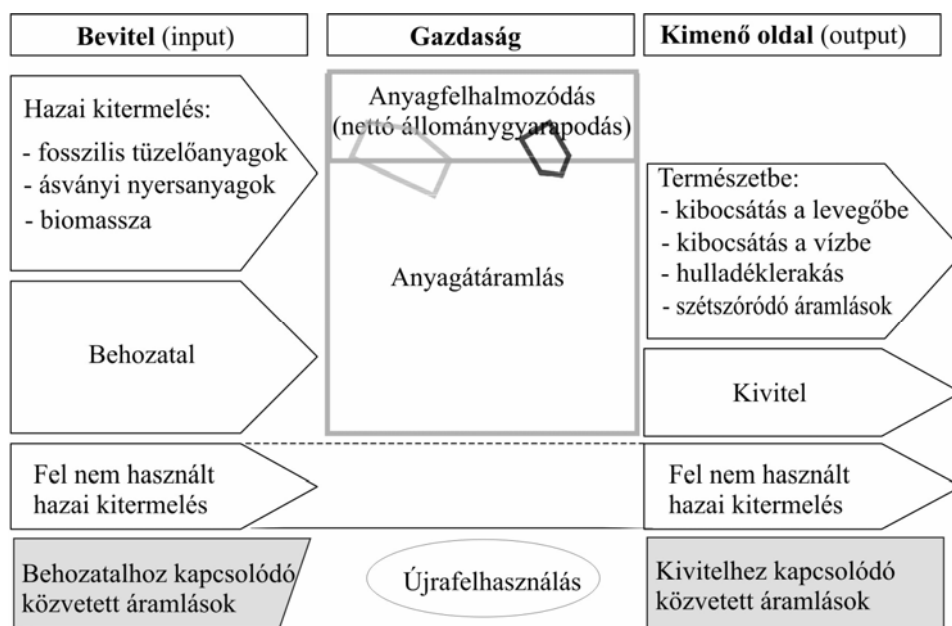
Az anyagáramlás-elemzések általában az „Ia”-, „Ib”-, „IIb”- és „IIc”-típusokat foglalják magukban, az „Ic”-típus inkább az életciklus-elemzés kategóriájába tartozik. A „IIa”-típus pedig inkább a környezetmenedzsmenthez köthető. Különböző kombinációi lehetnek a területi és termékorientált elemzéseknek. Közös jellemzője valamennyi elemzési típusnak az, hogy az anyaginput és -output folyamatok elszámolását kvantitatív módon alkalmazza a rendszerszemléletet követve.

4. Az anyagáramlások makrogazdasági modellje

Az anyagáramlások különböző osztályainak kialakításával és a fizikai állomány fontosságának hangsúlyozásával bevezethető az az általános mérlegmodell, amely

tartalmazza az összes lényeges beviteli és kimenő áramlást. A 4. ábrán vázolt rendszer feltárja a gazdaság fizikai anyagcseréjének alkotóelemeit, továbbá leírja a hazai kitermelést, behozatalt és kivitt fizikai egységekben, a gazdaság infrastruktúrájának fizikai növekedését, valamint a természetbe visszajuttatott anyagok mennyiségét.

4. ábra. A nemzetgazdasági MFA általános kerete (viz- és levegőáramlások nélkül)



Forrás: Eurostat [2001] nyomán.

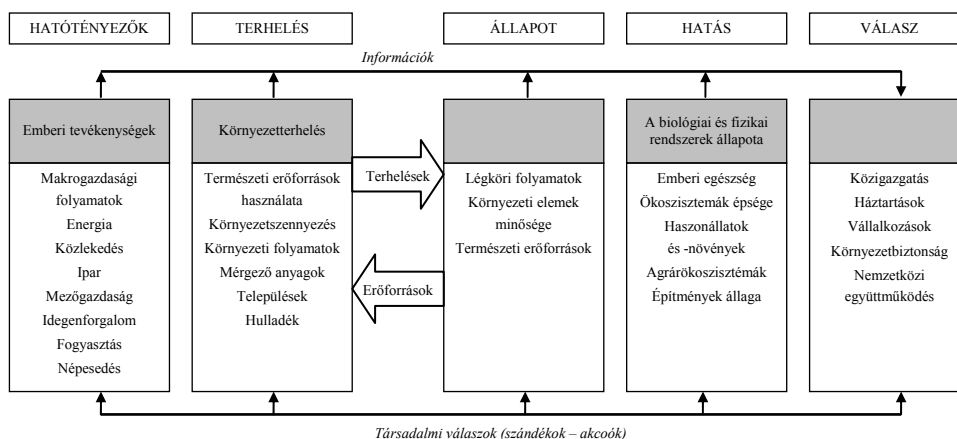
A gazdasági rendszerbe történő anyagbevétel magában foglalja a különféle anyagcsoportok (fosszilis tüzelőanyagok, ásványok és ércek, biomassza) felhasznált hazai (ki)termelését. Ezenkívül az anyagbevétel tartalmazza az úgynevezett fel nem használt hazai (ki)termelést (Unused Domestic Extraction – UDE) is. Az UDE megnevezés azokat az anyagokat foglalja magába, amelyeket meg kell mozgatni a kitermelő tevékenységek során, de amelyek nem lépnek be a gazdasági rendszerbe további feldolgozás céljából (például a bányászatból eredő takaró- és fedőrétegek, a mezőgazdasági betakarítás maradékai). Következésképpen a fel nem használt anyagok árama gazdaságilag nem jelenít meg értéket. Ezeket az áramlásokat a korábbi MFA-vizsgálatokban hazai és külföldi rejtett áramlásoknak (Domestic Hidden Flow – DHF és Foreign Hidden Flow – FHF) nevezték, minthogy láthatatlanok a pénzgazdaság számára (Adriansee et al. [1997]). Végül az anyagbevétel tartalmazza a fizikai behozatalt és az azzal kapcsolatos közvetlen áramlásokat is.

Az anyagbevitel során a gazdaságba jutó anyagok a társadalmi-gazdasági rendszerben vagy az állományt gyarapítják (nettó állománygyarapodás, például az infrastruktúra és tartós fogyasztási javak), vagy a hazai fogyasztásban jelennek meg az elszámolási időszakban (legtöbb esetben egy év), és így végső soron a rendszerhatáron keresztül hulladék és kibocsátások formájában a természetbe jutnak ki, vagy pedig más gazdaságokba kerülnek kivitelre.

5. Az anyagáramlásokból származtatott mutatók

A környezeti mutatók nemzetközi szinten összehangolt osztályozási rendszerében (az OECD *terhelés–állapot–válasz* kerete [1994], az Európai Unió *hatótényező–terhelés–állapot–hatás–válasz* kerete [1999]) az anyagáramlásokon alapuló mutatók a terhelésmutatók közé tartoznak. Ezek a mutatók azonosítják és leírják azokat a társadalmi-gazdasági tevékenységeket, amelyek különféle módon terhelik a környezetet. A környezet állapotát jellemző mutatók előállítása sok esetben nehéz vagy igen költséges, ezért a terhelésre vonatkozó mutatók gyakran használatosak az állapot közvetett mérésére. A tényleges környezeti hatásokra vonatkozó mutatók jelenleg csak meglehetősen korlátozott formában adnak információt. A társadalmi válaszok mutatóit a gyakorlatban elsősorban az utólagos szennyezéscsökkentési és -ellenőrzési intézkedések tartalmazzák, minthogy a megelőzésre és az intézkedéscsomagokra vonatkozó adatok előállítása sokkal nehezebb.

5. ábra. A környezeti mutatók fogalmi kerete



Forrás: Szabó E. – Pomázi I. [2003].

A nemzetgazdasági szintű anyagáramszámlákból több, az erőforrás-használatot jól jellemző mutató képezhető, amelyek három csoportba sorolhatók: 1. bevételmutatók, 2. a kimenőoldal mutatói és a 3. felhasználás- vagy fogyasztásmutatók.

A fő bevételmutatók a következők:

a) *Közvetlen anyagbevétel* (Direct Material Input – DMI): valamennyi gazdasági értékkel rendelkező és a termelésben vagy a fogyasztásban felhasznált anyagot magában foglalja. A hazai kitermelés és a behozatal összege.

b) *Összes anyagbevétel* (Total Material Input – TMI): a DMI és a fel nem használt hazai kitermelés összege.

c) *Összes anyagszükséglet* (Total Material Requirement – TMR): a TMI és behozattal kapcsolatos közvetett (felhasznált és fel nem használt) áramlásainak összege. A TMR így a legátfogóbb anyagbeviteli mutató, amely valamennyi beviteli áramlást tartalmazza.

$$\begin{aligned} \text{DMI} &= \text{DE} + \text{behozatal} \\ \text{TMI} &= \text{DMI} + \text{DHF} \\ \text{TMR} &= \text{TMI} + \text{FHF} = \text{DMI} + \text{DHF} + \text{FHF} \end{aligned}$$

A fő kimenőoldali mutatók a következők:

a) *Hazai (feldolgozásból származó) kibocsátás* (Domestic Processed Output – DPO): a természetbe történő kibocsátást jelenti, amely a hazai kitermelésből és behozatalból származó anyagok felhasználásához kapcsolódik. A DPO tartalmazza a levegőbe és a vízbe történő kibocsátásokat, a lerakókban elhelyezett hulladékot és a szét-szóródó áramlásokat. A DPO nem tartalmazza az újrafelhasznált anyagokat.

b) *Összes hazai kibocsátás* (Total Domestic Output – TDO): a DPO és a fel nem használt (ki)termelésből származó anyaglerakás összege. A mutató a (hazai) gazdasági tevékenységek nyomán a környezetbe kerülő összes anyagmennyiséget jelenti.

c) *Közvetlen anyagkibocsátás* (Domestic Material Output – DMO): a gazdaságot elhagyó anyagok összes mennyisége akár a hazai környezetbe történik a kibocsátás, akár árukivitel formájában külföldre kerül. A DPO és a kivitel összege.

d) *Összes anyagkibocsátás* (Total Material Output – TMO): a gazdaságot elhagyó anyagok összmenyisége, a TDO és kivitel összege.

DPO = kibocsátás + lerakás + szétszóródó áramlás

TDO = DPO + DHF

DMO = DPO + kivitel

TMO = TDO + kivitel

A fő felhasználásmutatók a következők:

a) *Hazai anyagfelhasználás* (Domestic Material Consumption – DMC): a gazdaságban felhasznált összes anyagot jelenti, kivéve a közvetett áramlásokat. Így a DMC a legközelebbi megfelelője a hagyományos nemzeti számlákban megjelenő összegzett bevételnek. A DMC a DMI kivittel csökkentett része.

b) *Összes anyagfelhasználás* (Total Material Consumption – TMC): a DMC és a behozattal és kivittel kapcsolatos közvetett áramlások összege, amely vagy életciklus-elemzési típusú vagy input–output módszerek felhasználásával számítható. A TMC a TMR kivittel és a hozzá kapcsolódó közvetett áramlásokkal csökkentett része.

c) *Fizikai kereskedelmi mérleg* (Physical Trade Balance – PTB): A PTB azt mutatja meg, hogy vajon a külföldről származó erőforrás-behozatal vagy az erőforrás-kivitel mértéke a nagyobb-e, és a hazai anyagfelhasználás mennyire alapul a hazai erőforrás-kitermelésen vagy a külföldről történő behozatalon. A fizikai kereskedelmi mérleg kétféle módon állítható össze. A közvetlen anyagáramlások PTB-je egyenlő az ország vagy régió behozatalának és kivitelének különbségével. Az átfogó PTB a behozattal és kivittel kapcsolatos közvetett áramlások figyelembevételével is kiszámolható.

d) *Nettó állománygyarapodás* (Net Addition to Stock – NAS): a gazdaságban évente felgyülemelő anyagmennyiséget mutatja, így a „gazdasági fizikai növekedése” kifejezés is használható. Az állományalkotó anyagok főleg az új infrastruktúrák építésére felhasznált építőanyagokat és a tartós fogyasztási cikkeket (például autók és gépek) jelentik.

$$\mathbf{DMC} = \mathbf{DMI} - \mathbf{kivitel}$$

$$\mathbf{TMC} = \mathbf{TMR} + \mathbf{behozatal} - \mathbf{kivitel} \pm \mathbf{FHF}$$

$$\mathbf{PTB} = \mathbf{behozatal} - \mathbf{kivitel}$$

$$\mathbf{NAS} = \mathbf{DMC} - \mathbf{DPO}$$

(A tanulmány második, befejező részét a *Statisztikai Szemle* következő számában közöljük.)

Tóth Zoltán

docens,

Károly Róbert Főiskola

E-mail: zoli@karolyrobert.hu

Nemlineáris függvények illesztésének néhány kérdése

A nemlineáris regressziós és trendfüggvények illesztésekor számos esetben alkalmazzuk az adatok transzformációjának módszerét. Ekkor az összefüggésre és az adatokra egy alkalmas transzformációt alkalmazva, paramétereinket lineáris összefüggésből kell meghatározni, amelynek normál egyenletrendszerét már könnyedén, akár zsebszámológép használatával is meg tudjuk oldani.

A probléma ismert a szakirodalomban, azonban a szerzők, kiváltképp a tankönyvek szerzői nem fordítanak kellő figyelmet arra, hogy a linearizálásból adódó torzítás milyen mértékű és irányú, hanem többnyire megelégszenek azzal a megállapítással, hogy a torzítás mértéke általában olyan, hogy az összefüggés a gyakorlat számára még használható (például *Hunyadi-Vita* [2004]). A tapasztalat azt mutatja, hogy ezek az eltérések exponenciális vagy hatványfüggvények illesztése esetén valóban nem mindig jelentősek, de például hiperbolikus függvény illesztésekor nagyságrendileg nagyobb eltéréseket is tapasztalhatunk, sőt ismeretes, hogy bonyolultabb függvények (például a logisztikus függvény) esetében ezek a torzítások gyakran használhatatlanná teszik a linearizálással kapott eredményeket. Az eltérések adatainktól függően egyazon függvénytípus esetén is eltérőek lehetnek.

Ennek a tanulmánynak az a célja, hogy az említett transzformációkból adódó torzításokat elemezze, és ezáltal is felhívja a figyelmet azokra a problémákra, melyek felett kiváltképp a tapasztalatlan alkalmazók hajlamosak átsiklani. Ennek megfelelően a kérdés felvetése után először egy egyszerűsített feladaton, a két, illetve több megfigyelés (számérték) legkisebb négyzetekkel kapható közepének meghatározásakor mutatjuk be az oda-vissza alkalmazott transzformációk tulajdonságait, majd a nemlineáris görbeillesztés néhány gyakori esetét vizsgáljuk meg. A tanulmányt a következtetések összefoglalása, a felhasználók számára megfogalmazott javaslatok, valamint a megválaszolatlanul maradt kérdések zárják.

1. A legkisebb négyzetek módszerének tulajdonságai az adatok transzformációja esetén

Amennyiben nemlineáris regressziós függvényt illesztünk, gyakran transzformáljuk a függő vagy a független változót. Az alapötlet első pillantásra helyesnek tűnik, hiszen tökéletesen illeszkedő függvények esetén az oda-vissza transzformáció (ha létezik) helyben hagyja a függvényt. Ha például egy pontosan hiperbolikus függvény szerint alakuló adatsorra akarunk függvényt illeszteni, akkor a függő változó reciprokára illesztve a lineáris függvényt, majd ennek a lineáris függvénynek a reciprokát véve megkapjuk az eredeti adatokra illeszkedő hiperbolikus függvényt.

Ez a gondolat annyira egyszerű, hogy nem is érdemes tovább magyarázni. A probléma ott kezdődik, amikor az illeszkedés nem tökéletes, hiszen akkor az eredmények korántsem lesznek ilyen triviálisak. Annak érdekében, hogy a problémát érthetően megvilágítsuk, először egy egyszerűbb feladatot fogalmazzunk és oldunk meg: azt vizsgáljuk meg, hogy két adat M , illetve m között keresve a legkisebb négyzetes eltérést teljesítő x értéket hogyan változik x értéke, ha az adatokat transzformáljuk, majd a transzformált adatokra kapott értéket az inverz transzformációval visszatranszformáljuk. Az előzők alapján nyilvánvaló, hogy amennyiben $M = m$, akkor x is egyenlő lesz velük, más esetekben azonban nem ez a helyzet. Az itt következő kis elemzés előtanulmányának is tekinthető a görbeillesztés feladathoz, ugyanakkor önállóan is érdekes, mivel rámutat egyes statisztikai átlagok tulajdonságaira.

1.1. Alapeset (nincs transzformáció)

Két érték, M és m ($M \neq m$) között keresünk egy olyan x értéket, amelyre

$$f(x) = (M-x)^2 + (m-x)^2 \rightarrow \min. \quad /1/$$

A függvény láthatóan felfelé nyíló parabola, ott lesz minimális, ahol $f'(x) = 0$. Mivel $f'(x) = -2(M-x) - 2(m-x) = 0$, ezt átrendezve

$$M - x + m - x = 0, \text{ majd } x = \frac{M + m}{2}$$

kapható, tehát a négyzetes eltérés akkor lesz a legkisebb, ha x a két érték számtani közepe. Az eredmény általánosan ismert alapösszefüggés, és több adat esetén is igaz.

1.2. A lineáris transzformáció esete

Legyen $l(x) = a + b \cdot x$ lineáris transzformáció. Ekkor $l(M)$ és $l(m)$ értékek között keressünk egy olyan y értéket, amelyre

$$f(y) = (l(M) - y)^2 + (l(m) - y)^2 \rightarrow \min, \quad /2/$$

majd ezt az y értéket az l inverz transzformációjával visszaalakítva az a kérdés, hogy visszakapjuk-e az alapesetben meghatározott x értéket.

Ekkor $f(y) = (a + b \cdot M - y)^2 + (a + b \cdot m - y)^2$, ami ott lesz minimális, ahol $f'(y) = 0$. Ekkor az

$$f'(y) = -2(a + b \cdot M - y) - 2(a + b \cdot m - y) = 0$$

egyenlet megoldását keressük, amire azt kapjuk, hogy

$$a + b \cdot M - y + a + b \cdot m - y = 0, \text{ majd } y = a + b \frac{M + m}{2}, \text{ végül}$$

$$l^{-1}(y) = -\frac{a}{b} + \frac{1}{b}y = \frac{M + m}{2}.$$

Ez azt jelenti, hogy a transzformált adatokra meghatározva a legkisebb négyzetes eltérést adó értéket, majd ezen értékre végrehajtva a transzformáció inverzét az adatok számtani átlagát kapjuk. Tehát transzformációval meghatározva y értékét, valamint inverz transzformációval x értékét az alapesetnek megfelelő x értéket kapjuk.

1.3. A reciprokképzés esete

Legyenek adva a M és m értékek. Legyen a transzformáció a reciprokképzés, amelynek inverze önmaga. Keressük azt az y értéket, amelyre:

$$f(y) = \left(\frac{1}{M} - y\right)^2 + \left(\frac{1}{m} - y\right)^2 \rightarrow \min. \quad /3/$$

A szokásos módon ezúttal is az $f'(y) = 0$ egyenletet kell megoldani:

$$f'(y) = -2\left(\frac{1}{M} - y\right) - 2\left(\frac{1}{m} - y\right) = 0, \quad y = \frac{\frac{1}{M} + \frac{1}{m}}{2}, \quad \text{azaz } x = \frac{1}{y} = \frac{2}{\frac{1}{M} + \frac{1}{m}}.$$

Ez azt jelenti, hogy a transzformált adatokra meghatározva a legkisebb négyzetes eltérést adó értéket, majd ezen értékre végrehajtva a transzformáció inverzét az adatok harmonikus átlagát kapjuk. Ez azzal a következménnyel jár, hogy az így kapott x érték nem egyezik meg az alapesetben meghatározott számtani átlaggal, hanem annál kisebb. Tehát a transzformált adatokra meghatározva a legkisebb négyzetek elvét teljesítő y értéket, majd azt az inverz transzformációval visszaalakítva, az eredeti adatok között a legkisebb négyzetek elvét teljesítő számtani átlagnál kisebb értéket, a harmonikus átlagot kapjuk eredményként.

1.4. A logaritmusképzés esete

Ismét legyen adva M és m és legyen a transzformáció a természetes alapú logaritmus, melynek inverze: e^x . Keressük azt az y értéket, amelyre:

$$f(y) = (\ln M - y)^2 + (\ln m - y)^2 \rightarrow \min. \quad /4/$$

Tudjuk, ez akkor teljesül, ha $f'(y) = -2(\ln M - y) - 2(\ln m - y) = 0$, ahonnan $y = \frac{\ln M + \ln m}{2}$, majd $x = e^y = e^{\frac{\ln M + \ln m}{2}} = \sqrt{M \cdot m}$ adódik.

Ekkor tehát a transzformált (logaritmizált) adatokra meghatározva a legkisebb négyzetek elvét teljesítő y értéket, majd azt az inverz transzformációval visszatranszformálva, az eredeti adatok között a legkisebb négyzetek elvét teljesítő számtani átlagnál kisebb értéket, a mértani átlagot kapjuk.

1.5. A négyzetgyökvonás esete

Legyen adva továbbra is M és m , és a transzformáció a négyzetgyökvonás, melynek inverze a négyzetre emelés. Keressük azt az y értéket, amelyre:

$$f(y) = (\sqrt{M} - y)^2 + (\sqrt{m} - y)^2 \rightarrow \min. \quad /5/$$

Tudjuk, ez akkor teljesül, ha $f'(y) = -2(\sqrt{M} - y) - 2(\sqrt{m} - y) = 0$, azaz

$$y = \frac{\sqrt{M} + \sqrt{m}}{2}, \text{ és } x = y^2 = \frac{M + m + 2\sqrt{M \cdot m}}{4}.$$

Az így kapott (visszatranszformált) x érték nem az adatok számtani átlaga, mert:

$$x = \frac{M + m + 2\sqrt{M \cdot m}}{4} \leq \frac{M + m}{2}, \text{ hiszen } M + m + 2\sqrt{M \cdot m} \leq 2(M + m) \text{ és } \sqrt{M \cdot m} \leq \frac{M + m}{2} \text{ a számtani és mértani átlag közötti ismert relációból adódóan.}$$

Láthatjuk, hogy a transzformált adatokra meghatározva a legkisebb négyzetek elvét teljesítő y értéket, majd azt az inverz transzformációval visszatranszformálva, egy a számtani átlagnál kisebb x értéket kapunk. Vagyis a visszatranszformált érték nem teljesíti a legkisebb négyzetek elvét az eredeti adatok között, hanem annál kisebb.

1.6. A négyzetre emelés esete

Legyen adva M és m , és legyen a transzformáció most a négyzetre emelés, melynek inverze a négyzetgyökvonás. Keressük azt az y értéket, amelyre:

$$f(y) = (M^2 - y)^2 + (m^2 - y)^2 \rightarrow \min. \quad /6/$$

Tudjuk, ez akkor teljesül, ha $f'(y) = -2(M^2 - y) - 2(m^2 - y) = 0$, azaz

$$y = \frac{M^2 + m^2}{2}, \text{ és } x = \sqrt{y} = \sqrt{\frac{M^2 + m^2}{2}}.$$

Az így kapott (visszatranszformált) x érték nem az adatok számtani átlaga, hanem annál nagyobb, mert:

$$\sqrt{\frac{M^2 + m^2}{2}} \geq \frac{M + m}{2}, \text{ hiszen}$$

$$\frac{M^2 + m^2}{2} \geq \frac{M^2 + m^2 + 2 \cdot M \cdot m}{4} \text{ és } \frac{M^2 + m^2}{2} \geq M \cdot m = \sqrt{M^2 \cdot m^2},$$

ami igaz az M^2 és m^2 számtani és mértani átlaga közötti ismert nagyságrendi összefüggés alapján. Ekkor tehát a transzformált adatokra meghatározva a legkisebb négyzetek elvét teljesítő y értéket, majd azt az inverz transzformációval visszatranszformálva, egy a számtani átlagnál nagyobb x értéket kapunk. Vagyis a visszatranszformált érték nem teljesíti a legkisebb négyzetek elvét az eredeti adatok között, hanem annál nagyobb.

1.7. Általános eset

Legyen $g(x)$ egy a vizsgált intervallumon monoton, folytonos és így invertálható függvény. Ekkor a $g(M)$ és $g(m)$ értékek között keresünk egy olyan y értéket, amelyre

$$f(y) = (g(M) - y)^2 + (g(m) - y)^2 \rightarrow \min. \quad /7/$$

$f(y)$ ott lesz minimális, ahol $f'(y) = -2(g(M) - y) - 2(g(m) - y) = 0$. Ekkor

$$y = \frac{g(M) + g(m)}{2}, \quad x = g^{-1}(y) = g^{-1}\left(\frac{g(M) + g(m)}{2}\right).$$

Ha $g^{-1}(y) = \frac{M + m}{2}$, akkor a legkisebb négyzetek elvének megfelelő x érték meghatározása és a g függvény által meghatározott transzformáció alkalmazási sorrendje felcserélhető. Ehhez a következő összefüggés kell, hogy teljesüljön:

$$g^{-1}\left(\frac{g(M) + g(m)}{2}\right) = \frac{M + m}{2}.$$

Az előző összefüggés pedig csak akkor teljesül, ha g inverze és így g is lineáris.

Összegezve megállapítható, hogy ha adatainkat transzformáljuk, csak a lineáris transzformáció hagyja változatlanul a két érték közötti legkisebb négyzetek elvét teljesítő értéket, azaz a számtani átlagot. Mivel a levezetések több adat esetén is ugyanúgy végigvihetők, megállapíthatjuk, hogy a transzformált adatok várható értékét meghatározva majd ezt az értéket az inverz transzformációval visszatranszformálva, csak lineáris transzformáció esetén kapjuk meg az eredeti adatok várható értékét.

Az eddig elmondottakhoz még egy megjegyzést kell fűznünk. Tudjuk, hogy ha g monoton, folytonos (nem lineáris) függvény és f folytonos, akkor amennyiben f -nek

létezik minimum helye, akkor $g(f)$ összetett függvénynek is létezik, és ez a két minimumhely egybeesik. Azonban esetünkben ezt a tulajdonságot nem alkalmazhatjuk, hiszen mi a következő típusú függvényeket akarjuk minimalizálni:

– az eredeti adatokra

$$F(a, b) = \sum_{i=1}^n (y_i - f(x_i))^2 \rightarrow \min ,$$

– a transzformált adatokra

$$G(a, b) = \sum_{i=1}^n (g(y_i) - g(f(x_i)))^2 \rightarrow \min .$$

Itt $G \neq g(F)$ és pontosan azt mutattuk meg az előző levezetésekben, hogy ez a minimum eltolódik a transzformációk hatására.

Korábban láttuk, hogy ha adataink illeszkednek, például egy hiperbolikus függvényre és azokat linearizáljuk, valamint a linearizált adatokra lineáris függvényt illesztünk, majd ezt az inverz transzformációval visszaalakítjuk, akkor visszkapjuk a kiindulási függvényt. Vagyis a transzformáció nem torzítja az eredményt. Ennek oka, hogy tökéletes illeszkedés esetén ezt a gondolatmenetet követve olyan esettel állunk szemben, ahol $M = m$, és így ekkor x értéke is illeszkedik $M = m$ értékre. Feladatainkban ilyen ideális eset nincs is, mert ekkor transzformáció nélkül is meg tudnánk találni a függvényt, illetve a paraméterek értékét. Viszont ha az adatok nem illeszkednek pontosan az eredeti vagy a transzformált függvényre, akkor a transzformáció eltérít minket a legjobban illeszkedő függvénytől.

A kétszeres transzformáció természetesen nem biztosítja automatikusan azt, hogy a kiinduló és a transzformációk után kapott értékek azonosak legyenek. Egy egyszerű példát említve illeszkedjenek az eredeti adatpárok egy egyenesre. A transzformáció legyen a reciprok képzése, és a transzformált adatokra (amelyek nyilván nem illeszkednek tökéletesen egy egyenesre) illesztünk lineáris függvényt, majd a transzformált adatokra kapott lineáris függvényt és annak értékeit alakítsuk vissza az inverz transzformációval. Az így kapott visszatranszformált adatok nyilvánvalóan nem egyeznek az eredetiekkel.

2. A hibatagok változása a transzformáció hatására

Vizsgáljuk meg a továbbiakban, hogy milyen mennyiségi összefüggés van az eredeti függvényre illesztett nemlineáris regressziós vagy trendfüggvény hibatagjai és a linearizált függvény hibatagjai között. Bevezetőül meg kívánjuk jegyezni, hogy

a transzformáció mindig az y függvényértékekre vonatkozik, illetve, hogy mindig additív hibagról fogunk beszélni. Az eredeti függvény hibáját az i -edik adat esetén ε_i , a transzformált függvény esetén ugyanezt δ_i jelöli.

Az itt következő elemzésekben meghatározzuk néhány alapvető nemlineáris esetben a hibatagok négyzetösszegét és keressük ezen összegek minimumhelyeit. Megmutatjuk, hogy ezen minimumhelyek az adattranszformáció alkalmazásával végzett megoldásokban és az eredeti adatokra vizsgálva nem esnek egybe. Egyben azt is igyekszünk megmutatni, hogy mi a viszony a kétféle hibatag között.

2.1. A hibatagok viszonya exponenciális függvény esetén

Ebben az esetben az illesztett függvény $y = a \cdot e^{bx}$ alakú, a transzformáció pedig, amivel ez linearizálható, a természetes alapú logaritmus. Megjegyezzük, hogy leggyakrabban az $a > 0$, $b > 0$ esettel találkozunk, így példánkban is ilyet mutattunk be.

Természetesen az itteni elemzés kiterjeszhető a függvény más parametrizálására is. Az eredeti adatokra, illetve az exponenciális függvényre vonatkozó additív hibatag legyen ε , a transzformált adatokra lineáris függvényt illesztünk és az itt szereplő additív hibatag legyen δ .

Az a és b paraméterek változtatásával eredeti célunk a $\sum \varepsilon_i^2 \rightarrow \min$ meghatározása. Ennek egyik útja volt az adatok transzformálása és $\sum \delta_i^2 \rightarrow \min$ meghatározása. A következőkben bemutatjuk, hogy ez a két minimum nem esik egybe, és meghatározzuk, hogy mekkora az eltérés az ε és δ értékek között. Legyen az eredeti adatokra illesztendő függvény alakja $y_i = a \cdot e^{b \cdot x_i} + \varepsilon_i$, amiből azonnal kapható az $y_i - \varepsilon_i = a \cdot e^{b \cdot x_i}$ forma. A transzformált adatokra illesztett lineáris függvény:

$$\ln y_i = \ln a + b \cdot x_i + \delta_i, \text{ ahonnan } \ln y_i - \delta_i = \ln a + b \cdot x_i, \text{ és } e^{\ln y_i - \delta_i} = \frac{y_i}{e^{\delta_i}} = a \cdot e^{b \cdot x_i}.$$

A két utóbbi kapott alak összevetése és némi átalakítás után a következő adódik:

$$y_i - \varepsilon_i = \frac{y_i}{e^{\delta_i}}; \frac{1}{e^{\delta_i}} = \frac{y_i - \varepsilon_i}{y_i}; \frac{1}{e^{\delta_i}} = e^{-\delta_i} = \left(1 - \frac{\varepsilon_i}{y_i}\right) \text{ és innen}$$

$$-\delta_i = \ln \left(1 - \frac{\varepsilon_i}{y_i}\right) = \ln \left(1 - \frac{1}{\frac{y_i}{\varepsilon_i}}\right) = \frac{\varepsilon_i}{y_i} \ln \left(1 - \frac{1}{\frac{y_i}{\varepsilon_i}}\right).$$

Mivel $\lim_{x \rightarrow \infty} \left(1 - \frac{1}{x}\right)^x = e^{-1}$ ezért nagy y_i és kicsi ε_i , tehát jó közelítés esetén eb-

ből következik, hogy $\ln \left(1 - \frac{1}{\frac{y_i}{\varepsilon_i}}\right)^{\frac{y_i}{\varepsilon_i}} \approx -1$, ami miatt

$$\delta_i \approx \frac{\varepsilon_i}{y_i}, \text{ azaz } \varepsilon_i \approx y_i \delta_i. \quad /8/$$

Látható, hogy $\sum \varepsilon_i^2$ és $\sum \delta_i^2$ minimuma nem feltétlenül egyezik meg. Pontosabban, mivel $\sum \frac{\varepsilon_i^2}{y_i^2} \approx \sum \delta_i^2$ és mi $\sum \delta_i^2 \rightarrow$ minimumát határozzuk meg, ezért az ez-

zel egyenlő $\sum \frac{\varepsilon_i^2}{y_i^2}$ összegben a nagy y értékekhez tartozó, ε_i hibák kisebb súllyal

szerepelnek. Ez azt jelenti, hogy a transzformációval kapott regressziós függvényünk nagy y értékekre nem lesz annyira pontos, vagyis a transzformációval valóban nem kapjuk meg a legjobban illeszkedő exponenciális függvényt. A legjobban illeszkedő függvény meghatározásához a $\sum \varepsilon_i^2 \rightarrow$ minimumot kellett volna meghatározni, ami viszont analitikusan nem kezelhető normálegenlet-rendszerhez vezet.

Alakítsuk most vissza regressziós függvényünk transzformált alakját hibataggal együtt:

$$\begin{aligned} \ln y_i &= \ln a + b \cdot x_i + \delta_i, \\ y_i &= a \cdot e^{b \cdot x_i} \cdot e^{\delta_i} = a \cdot e^{b \cdot x_i} \left(1 + \frac{P_i}{100}\right). \end{aligned}$$

Ekkor az látható, hogy az előző felírás esetén $e^{\delta_i} = 1 + \frac{P_i}{100}$ jelöléssel P_i azt mutatja meg, hogy hány százalék az eltérés az eredeti és számított y_i érték között, és mi ezt minimalizáljuk. Ez akár egy kritériuma is lehet a függvényillesztésnek (az eltérés százalékos összegének a minimalizálása), de jól látható, hogy ez nem felel meg a legkisebb négyzetek elvének.

Nézzük meg, hogy ez az eredmény hogyan látható a következő fiktív adatsoron:

1. példa

Legyen adva az alábbi 4 adatpár

Eredeti adatok				Transzformált adatok			$\ln\left(1 - \frac{\varepsilon_i}{y_i}\right)$
x_i	y_i	y_i számított (transzformáció)	ε_i	$\ln y_i$	$\ln y_i$ számított	δ_i	
0	2,0	1,379	0,621	0,693	0,321	0,372	-0,372
1	1,7	3,179	-1,479	0,531	1,156	-0,626	0,626
2	8,4	7,327	1,073	2,128	1,992	0,137	-0,137
3	19,0	16,889	2,111	2,944	2,827	0,118	-0,118

Ezen adatokat transzformálva és a transzformált adatokra alkalmazva a legkisebb négyzetek elvét (lineáris függvényt illesztve), majd a paramétereket visszatranszformálva $a = 1,379$, illetve $b = 0,8351$ értékeket kapunk, azaz a legjobb függvénynek a linearizálás módszerével az $y = 1,379 \cdot e^{0,8351x}$ függvény bizonyult. Ezen paraméterekkel kiszámítva a függvényértékeket, az előző ε_i , δ_i értékeket kapjuk, amelyek között megfigyelhetjük a levezetett összefüggést, miszerint $\varepsilon_i \approx y_i \delta_i$, illetve $-\delta_i = \ln\left(1 - \frac{\varepsilon_i}{y_i}\right)$. Ha ezután az y_i adatokra úgy illesztünk legjobban illeszkedő exponenciális függvényt, hogy nem végezzük el a linearizáló transzformációt, hanem az eredeti forma maradékainak négyzetösszegét minimalizáljuk (numerikus közelítéssel, iterációval), akkor a legjobban közelítő függvénynek az $y = 1,233 \cdot e^{0,914x}$ függvény bizonyul.

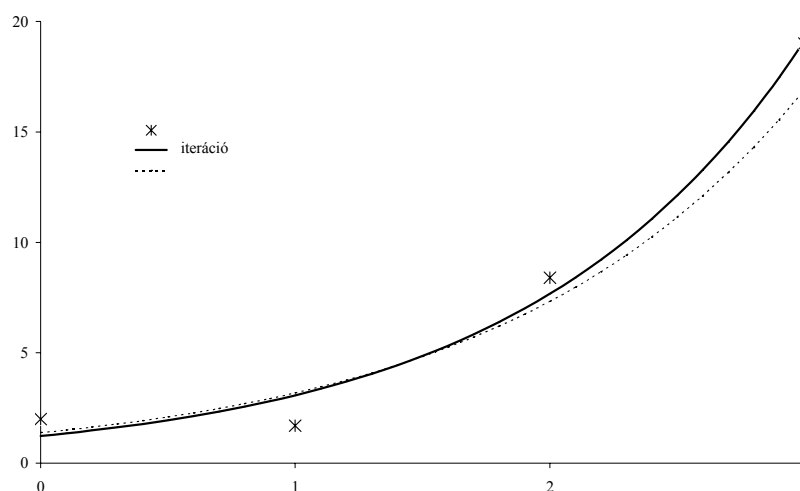
Az illeszkedések összehasonlítása érdekében kiszámítottuk az $I^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$ korrelációs indexet mind az iterációval meghatározott függvényre, mind a transzformációval meghatározott függvényre. Ekkor azt kaptuk, hogy:

$$- \text{transzformáció esetén: } y = 1,379 \cdot e^{0,8351x} \quad I^2 = 0,9584,$$

$$- \text{iteráció esetén: } y = 1,233 \cdot e^{0,914x} \quad I^2 = 0,9846.$$

Természetesen I^2 értékeiből is látszik, hogy a transzformált adatokra alkalmazva a legkisebb négyzetek elvét, majd az inverz transzformációval visszatranszformálva a paramétereket nem a legjobb regressziós függvényt kapjuk.

1. ábra. A transzformációval, illetve iterációval meghatározott exponenciális függvények



Az 1. ábrán is láthatók a levezetett összefüggések, miszerint a transzformációval meghatározott függvény „kisebb súllyal figyel” a nagyobb függvényértékek esetén elkövetett hibákra. Megjegyzendő, hogy a hatványfüggvény esetén hasonló nagyságrendű eltérést kapunk a transzformáció hatására.

Az itt tárgyalt exponenciális függvény egyik speciális esete az a konkáv függvény, amelyet leggyakrabban az $y = A(1 - ae^{bx})$, ($b < 0$) alakban szoktunk specifikálni, és amelyet egyszerű telítődési folyamatok leírására használhatunk. Az eddigiek alapján aligha meglepő az az állítás, miszerint a transzformált alakra alkalmazott legkisebb négyzetes illesztés ez esetben is más eredményt ad, mint az eredeti formára iterációs úton készített becslés.

Az eltéréseket illetően könnyen belátható, hogy a /8/-ban szereplő $\delta_i \approx \frac{\varepsilon_i}{y_i}$ össze-

függést ez esetben a $\delta_i \approx \frac{\varepsilon_i}{A - y_i}$ váltja fel, és a paraméterbecsléshez ennek négyzet-

összegét kell minimalizálni. Ennek következménye az, hogy a négyzetösszegben a kis y_i értékekhez tartozó ε_i hibák szerepelnek kisebb súllyal. Ez pedig azt jelenti, hogy a transzformációval kapott függvényünk a kis y értékekre lesz kevésbé pontos. Emlékeztetünk arra, hogy az eredetileg felírt (konvex) exponenciális függvény esetében ez fordítva volt. Ennek az eltérő eredménynek az egyszerű intuitív magyarázata az, hogy a telítődési szinthez közeledve egyre kisebb mozgástere van az illesztett görbének; egyre inkább, egyre kisebb ingadozásokkal simul a telítődési szinthez.

2.2. A hibatagok viszonya hiperbolikus regressziós függvény esetén

Ebben az esetben az illesztett függvény alakja: $y = \frac{1}{a + b \cdot x}$, a transzformáció a reciprokképzés, amelynek az inverze is a reciprokképzés lesz. Az eredeti adatokra és a hiperbolikus függvényre vonatkozó additív hibatag legyen ε , a transzformált adatokra lineáris függvényt illesztünk és az itt szereplő additív hibatag legyen δ .

Az a és b paraméterek változtatásával eredeti célunk a $\sum \varepsilon_i^2 \rightarrow \min$ meghatározása. Ennek egyik útja volt az adatok transzformálása és $\sum \delta_i^2 \rightarrow \min$ meghatározása. A következőkben bemutatjuk, hogy ez a két minimum nem esik egybe és meghatározzuk, hogy mekkora az eltérés az ε és δ értékek között.

Az eredeti adatokra illesztett függvény $y_i = \frac{1}{a + b \cdot x_i} + \varepsilon_i$, amit átalakítva

$$\frac{1}{y_i - \varepsilon_i} = a + b \cdot x_i \quad /9/$$

kapható. A transzformált alakból $\frac{1}{y_i} = a + b \cdot x_i + \delta_i$, majd

$$\frac{1}{y_i} - \delta_i = \frac{1 - y_i \cdot \delta_i}{y_i} = a + b \cdot x_i \quad /10/$$

adódik. A /9/ és /10/ jobb oldalának egyezősége miatt

$$\frac{1}{y_i - \varepsilon_i} = \frac{1 - y_i \cdot \delta_i}{y_i} \Rightarrow y_i = y_i - \varepsilon_i - y_i^2 \cdot \delta_i + y_i \cdot \delta_i \cdot \varepsilon_i,$$

amit átrendezve az $\varepsilon_i = \delta_i \cdot y_i^2 \cdot \left(\frac{\varepsilon_i}{y_i} - 1 \right)$ egyenlőséghez jutunk. Ekkor, ha $\left| \frac{\varepsilon_i}{y_i} \right|$ kicsi,

vagyis jó közelítés esetén $\frac{\varepsilon_i}{y_i} - 1 \approx -1$, így $\varepsilon_i \approx -\delta_i \cdot y_i^2$ azaz

$$\sum \varepsilon_i^2 \approx \sum \delta_i^2 \cdot y_i^4. \quad /11/$$

A korábbiakhoz hasonlóan itt is látható, hogy $\sum \varepsilon_i^2$ és $\sum \delta_i^2$ minimuma nem feltétlenül egyezik meg (sőt az egyezésnek kicsi az esélye), mivel $\sum \frac{\varepsilon_i^2}{y_i^4} \approx \sum \delta_i^2$ és mi $\sum \delta_i^2 \rightarrow$ minimumát határozzuk meg, ezért az ezzel egyenlő $\sum \frac{\varepsilon_i^2}{y_i^4}$ összegben a nagy y értékekhez tartozó ε_i hibák kisebb súllyal szerepelnek. Ez azt jelenti, hogy a transzformációval kapott regressziós függvényünk nagy y értékekre még kevésbé lesz pontos, mint exponenciális függvény esetén, vagyis a transzformációval továbbra sem kapjuk meg a legjobban illeszkedő hiperbolikus függvényt. A legjobban illeszkedő függvény meghatározásához a $\sum \varepsilon_i^2 \rightarrow$ minimumot kellett volna meghatározni. Nézzük meg, hogyan láthatók az eredmények egy egyszerű számpéldán.

2. példa

Tegyük fel, hogy az eredeti adatok körülbelül 10-20 százalékos hibával illeszkednek a normál hiperbolára

Eredeti adatok				Transzformált adatok			$\delta_i \cdot y_i^2 \left(\frac{\varepsilon_i}{y_i} - 1 \right)$
x_i	y_i	y_i számított	ε_i	$1/y_i$	$1/y_i$ szá- mított	δ_i	
0,1	12	2,5	9,5	0,083	0,4	-0,317	9,5
1	0,8	1,11	-0,31	1,25	0,902	0,349	-0,31
10	0,17	0,169	0,001	5,882	5,915	-0,033	0,001

Az értékeket transzformálva és a transzformált adatokra alkalmazva a legkisebb négyzetek elvét $a=0,3445$, illetve $b=0,557$ értékeket kapunk, azaz a legjobb függvénynek a linearizálás módszerével az $y = \frac{1}{0,3445 + 0,557 \cdot x}$ függvény bizonyul. Az ezen paraméterekkel kiszámított függvényértékeket és a hozzájuk tartozó ε_i , δ_i értékeket a 2. példa tartalmazza, amelyek között megfigyelhetjük a levezetett összefüggéseket, miszerint

$$\varepsilon_i \approx -\delta_i \cdot y_i^2, \text{ illetve } \varepsilon_i = \delta_i \cdot y_i^2 \cdot \left(\frac{\varepsilon_i}{y_i} - 1 \right).$$

Ha ezután az eredeti y_i adatokra iterációs módon illesztünk legjobban illeszkedő hiperbolikus függvényt az Excel-program segítségével, akkor a legjobban közelítő függvénynek az $y = \frac{1}{-0,045 + 1,28 \cdot x}$ függvény bizonyul.

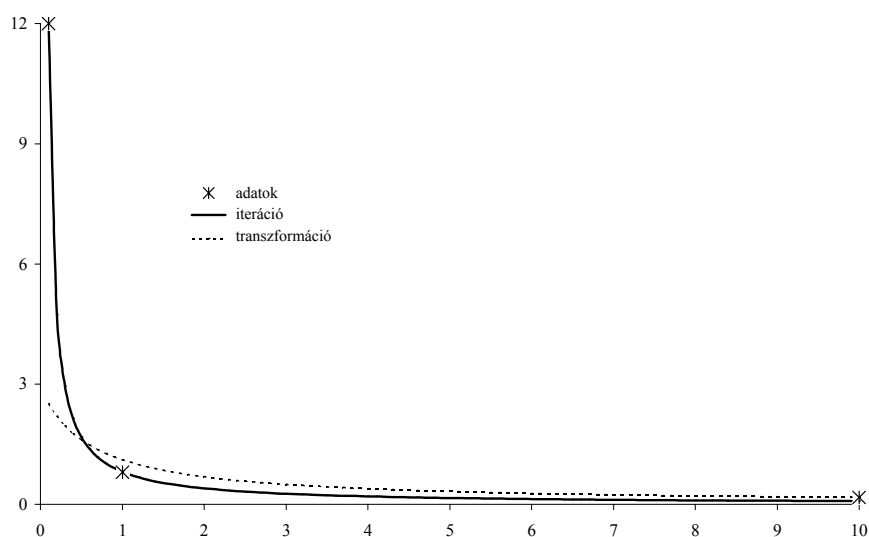
Az 1. példához hasonlóan itt is kiszámítottuk a korrelációs indexet mind az iterációval, mind pedig a transzformációval meghatározott függvényre:

$$\text{– transzformáció esetén: } y = \frac{1}{0,3445 + 0,557 \cdot x} \quad I^2 = -0,02 ,$$

$$\text{– iteráció esetén: } y = \frac{1}{-0,045 + 1,28 \cdot x} \quad I^2 = 0,9999 .$$

Annak felismerését, hogy nagy baj lehet az adatok transzformációjával történő regressziós függvényillesztéssel egy, a 2. példához hasonló feladat sugallta. Itt I^2 értéke negatív, ami azt jelenti, hogy a transzformációval meghatározott regressziós függvényünk rosszabb közelítő függvénye adatainknak, mint az \bar{y} konstans függvény. Ez alapján már várható volt, hogy a transzformációval meghatározott regressziós függvény ilyen esetben jelentősen eltér a legjobban illeszkedő regressziós függvénytől. Lehet olyan adatokat találni, ahol hasonló eredményt kapunk exponenciális függvény illesztése esetén is.

2. ábra. A transzformációval, illetve iterációval meghatározott hiperbolikus függvények



Az I^2 értéke alapján nyilvánvaló, hogy közelítőleg sem kaptuk meg eredményként a legjobb hiperbolikus regressziós függvényt. Az iterációs módszerrel pedig meg is határoztuk a legjobb hiperbolikus közelítő függvényt. A 2. példából az is látszik, hogy bár a δ_i értékek tűrhetőek, a nagy eredeti y érték esetén az $\varepsilon_i = 9,5$ tehát nagyon nagy, ahogy az előző levezetés alapján vártuk, hiszen $|\varepsilon| \approx |\delta| \cdot y^2$. Ezek az eredmények láthatók a 2. ábrán.

2.3. A hibatagok viszonya logisztikus regressziós függvény esetén

Ebben az esetben az illesztett függvény $y = \frac{A}{1 + a \cdot e^{b \cdot x}}$ alakú, az y érték transzformációja pedig összetett: A -val való osztás \rightarrow reciprok $\rightarrow -1 \rightarrow$ logaritmizálás.

Az eredeti és transzformált adatokra vonatkozó hibatagok jelölése ugyanaz mint eddig, és eredeti célunk itt is a $\sum \varepsilon_i^2 \rightarrow \min$ meghatározása. Ehelyett transzformáljuk az adatokat és a transzformáció utáni additív hibatagot minimalizáljuk, $\sum \delta_i^2 \rightarrow \min$. A következőkben itt is bemutatjuk, hogy ez a két minimum nem esik egybe és meghatározzuk, hogy mekkora az eltérés az ε és δ értékek között.

Az eredeti adatokra illesztett függvény $y_i = \frac{A}{1 + a \cdot e^{b \cdot x_i}} + \varepsilon_i$, átalakítva

$$\frac{A}{y_i - \varepsilon_i} - 1 = a \cdot e^{b \cdot x_i} \quad /12/$$

Az adatok transzformációja és a transzformált adatokra illesztett lineáris függvény leírása az alábbi: $\ln\left(\frac{A}{y_i} - 1\right) = \ln a + b \cdot x_i + \delta_i$ ezt átalakítva $\frac{A}{y_i} - 1 = a \cdot e^{b \cdot x_i} \cdot e^{\delta_i}$ majd

$$\left(\frac{A}{y_i} - 1\right) \cdot \frac{1}{e^{\delta_i}} = a \cdot e^{b \cdot x_i} \quad /13/$$

adódik. A /12/ és /13/ egyenlőségekből

$$\frac{A}{y_i - \varepsilon_i} - 1 = \left(\frac{A}{y_i} - 1\right) \cdot \frac{1}{e^{\delta_i}},$$

amit átrendezve

$$\frac{A - y_i + \varepsilon_i}{y_i - \varepsilon_i} = \frac{A - y_i}{y_i \cdot e^{\delta_i}}, \text{ majd innen}$$

$$y_i \cdot e^{\delta_i} = y_i - \frac{A \cdot \varepsilon_i}{A - y_i + \varepsilon_i}.$$

Tovább alakítva ezt

$$\frac{y_i}{A} (A - y_i + \varepsilon_i) \cdot (1 - e^{\delta_i}) = \varepsilon_i \text{ és mivel } \varepsilon_i \ll A - y_i, \text{ ezért}$$

$$(\varepsilon_i) \approx \left(\frac{y_i}{A} (A - y_i) \right) \cdot (1 - e^{\delta_i}), \text{ ahonnan}$$

$$(1 - e^{\delta_i})^2 \approx \frac{(\varepsilon_i)^2}{\left[-\frac{1}{A} \left(y_i - \frac{A}{2} \right)^2 + \frac{A}{4} \right]^2}.$$

Mivel δ_i^2 és $(1 - e^{\delta_i})^2$ minimuma egybeesik, $\sum \delta_i^2$ itt is ε_i^2 -k súlyozott összege és a súlyok láthatóan a szögletes zárójelben levő lefelé nyíló parabola értékeinek reciprokai. A legkisebb súly a nevező maximumához tartozik, ami azt jelenti, hogy az $y_i = \frac{A}{2}$ értékhez tartozó ε_i hiba szerepel a legkisebb súllyal míg 0-hoz illetve A -hoz közeli y_i értékek esetén az ε_i hiba nagyobb súllyal szerepel, ezért itt pontosabb lesz a transzformációval megalkotott regressziós függvény. Ez azt is jelenti, hogy a transzformációval kapott regressziós függvényünk $y_i = \frac{A}{2}$ értékekre lesz leginkább pontatlan.

3. példa

Illesszünk logisztikus regressziós függvényt az adatokra mind a transzformáció mind az iteráció módszerével. Legyen ebben a példában $A=100$. (A példára vonatkozó adatok a következő táblázatban találhatók.)

Eredeti adatok				Transzformált adatok			$\left[-\frac{1}{A} \left(y_i - \frac{A}{2} \right)^2 + \frac{A}{4} \right] \cdot (1 - e^{\delta_i})$
x_i	y_i	y_i számított (transzformáció)	ε_i	$\ln\left(\frac{A}{y_i} - 1\right)$	$\ln\left(\frac{A}{y_i} - 1\right)$ számított	δ_i	
0	3	4,407	-1,407	3,476	3,077	0,399	-1,427
1	20	13,949	6,051	1,386	1,820	-0,433	5,625
2	22	36,308	-14,310	1,265	0,562	0,704	-17,523
3	80	66,719	13,281	-1,386	-0,696	-0,691	7,981
4	95	87,577	7,423	-2,944	-1,953	-0,991	2,988
5	90	96,123	-6,123	-2,197	-3,211	1,013	-15,791

A kiindulási értékekre alkalmazva a leírt transzformációt és a transzformált adatokra lineáris függvényt illesztve $a = 21,693$, illetve $b = -1,258$ értékeket kaptunk, azaz a linearizálás módszerével adatainkra a legjobb logisztikus függvénynek az

$y = \frac{100}{1 + 21,693 \cdot e^{-1,258 \cdot x}}$ függvény bizonyult. Az adott paraméterértékeknek megfelelő

függvényre vonatkozó számított függvényértékek és a hozzájuk tartozó additív hibataragok szintén a táblázatban találhatóak. Azonban a kétszeres elhanyagolás miatt az

$\sum \frac{\varepsilon_i^2}{(A - y_i)^2} \approx \sum (1 - e^{\delta_i})^2 \approx \sum \delta_i^2$ összefüggésben, az ε_i értékeket csak kevésbé

pontosan követik az $(A - y_i) \cdot (1 - e^{\delta_i})$ értékek.

Ha ezután az eredeti adatokra iterációs módszerrel illesztünk legjobban illeszkedő

logisztikus függvényt, akkor ez a következő lesz $y = \frac{100}{1 + 139,66 \cdot e^{-2,028 \cdot x}}$.

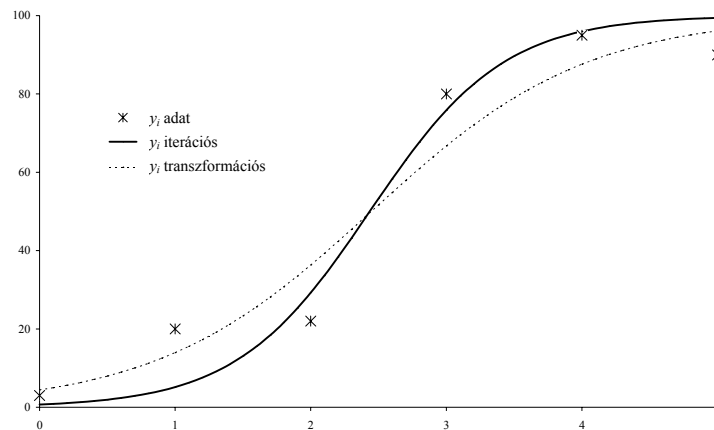
Mindkét függvény esetén kiszámítva a korrelációs indexet itt is érzékelhető az eltérés:

$$\text{– transzformáció esetén: } y = \frac{100}{1 + 21,693 \cdot e^{-1,258 \cdot x}} \quad I^2 = 0,939,$$

$$\text{– iteráció esetén: } y = \frac{100}{1 + 139,66 \cdot e^{-2,028 \cdot x}} \quad I^2 = 0,954.$$

A grafikonról látható, hogy jelentős eltérések vannak az iterációval meghatározott legjobban közelítő és a transzformációval meghatározott logisztikus függvényeink között. Az is látható a grafikonon, hogy a két középső érték esetén kapjuk a legnagyobb ε értéket.

3. ábra. A transzformációval, illetve iterációval meghatározott logisztikus függvények



3. Következtetések és nyitott kérdések

A korábbiakban bemutattuk azt, hogy a tankönyvek által javasolt linearizálás az esetek jó részében korántsem olyan ártatlan, mint ahogy azt olykor sugallják, hiszen ez a módszer a görbeillesztéskor nemritkán igencsak félrevezető eredményeket ad. Eredményeink nyilvánvalóan azt javasolják, hogy a linearizálás helyett (vagy inkább mellett) használjuk a nemlineáris legkisebb négyzetek módszerét, amely az eredeti nemlineáris feladatot oldja meg numerikus közelítéssel. Erre a gyorsan fejlődő számítástechnika egyre jobb lehetőségeket kínál.

Mint tudjuk, a gyakorlatilag bárki számára hozzáférhető Excel-programban lehetőségünk van a Solver-segédprogrammal szélsőérték meghatározására feltételek mellett és anélkül. A segédprogram használatakor meg kell adni, hogy melyik cella minimumát vagy maximumát keressük, valamint a szélsőérték típusát, és jelezni kell, hogy mely cellák iteratív változtatása mellett történjen a minimum meghatározása. Amennyiben vannak feltételek azokat is közölni kell.

Amikor regressziós függvények paramétereit kutatjuk, akkor a célcella minimumát keressük és a célcella tartalma a következő képlet és annak értéke:

$$\sum (y_i - y_{i, \text{számított}})^2 \rightarrow \min .$$

A módosuló cellák a paramétereket tartalmazzák és feltételek megadása nem szükséges. Amikor jelen írásban iterációs paraméter kereséséről beszélünk, mindig erre a módszerre gondolunk.

Az ezzel a módszerrel történő regressziós függvény meghatározásnál a paramétereknek kezdeti értéket kell ugyan adni, de elvben bármely függvénytípus esetén elkészíthető mechanikusan a paraméterek és a függvényértékek becslése. Amennyiben az eljárás valóban az abszolút minimumot eredményezi, a kapott paraméterek és függvényértékek torzítatlanok lesznek. Ugyanakkor felmerülhet az a kérdés, hogy a módszer minden függvénytípus és bármilyen kezdeti érték megadása esetén is konvergens-e, és ha igen a legkisebb négyzetekkel definiált függvény abszolút minimumát adja-e, vagy valamilyen lokális minimumot talál. Ezért érdemes azt a kérdést is felvetni, hogy bizonyos függvénytípusok esetében lehet-e analitikusan megoldható normálegyenleteket számolni. Végül megemlítjük, hogy amennyiben lehetőség van rá, érdemes egy feladatot több különböző, feltehetően eltérő megoldó algoritmusokat használó programcsomagok segítségével elvégezni. Az esetleg eltérő eredmények figyelmeztethetnek az említett hibákra.

Irodalom

- ALMÁSI A. – TÓTH Z. – TÖRCSVÁRI ZS. [2005]: *Have you ever seen a best exponential regression function?* Tagungsband. 9. Thüringisch – Ungarisches Symposium. Fachhochschule. Jena.
- HUNYADI L. – VITA L. [2004]: *Statisztika közgazdászoknak*. KSH. Budapest.
- MUNDRUCZÓ GY. [1981]: *Alkalmazott regressziószámítás*. Akadémiai Kiadó. Budapest.
- RAPPAI G. [2001]: *Üzleti statisztika Excellel*. KSH. Budapest.
- SZÜCS I. [2002]: *Alkalmazott statisztika*. Agroinform Kiadó. Budapest.

Hírek, események

**Vezetői megbízás visszavonása, cím-
adományozás.** *Dr. Pukli Péter*, a Központi Statisztikai Hivatal elnöke *Gárdos Évától*, 2006. február 14-ei hatállyal a Népesedés-, egészségügyi és szociális statisztikai főosztály vezetésére adott főosztályvezetői megbízását visszavonta, egyidejűleg szakmai munkája elismeréséül részére szakmai főtanácsadói címet adományozott.

Jutalom. Közszolgálati jogviszonyban töltött idejük alapján 2006. február hónapban jubileumi jutalomban részesült: 25 éves szolgálatért: *Simonné Józsa Edit*, Iparstatisztikai főosztály; 30 éves szolgálatért: *Steidl Károlyné*, Külkereskedelem-statisztikai főosztály; *Waffenschmidt Jánosné*, Adatgyűjtő főosztály; 35 éves szolgálatért: *Kovács Attila*, Adatgyűjtő főosztály; *dr. Lindnerné dr. Eperjesi Erzsébet*, Életszínvonal- és emberierőforrás-statisztikai főosztály; *Szabó Miklósné*, Adatgyűjtő főosztály; 40 éves szolgálatért: *dr. Herczeg András*, Iparstatisztikai főosztály; *Heves Józsefné*, Népesedés-egészségügyi és szociális statisztikai főosztály; *dr. Pukli Péter*, Elnökség; *dr. Rémes Jenőné*, KSH Győri Igazgatóság.

Éves munkaértekezlet. A Központi Statisztikai Hivatal 2006. február 9. és 10. között tartotta éves munkaértekezletét a Corinthia Aquincum Hotelben, illetve a Hivatal épületében. Az értekezletet *dr. Bagó Eszter*, a KSH elnökhelyettese nyitotta meg a Hivatal 2005. évi szakmai munkájáról szóló beszámolójával. Kiemelte, hogy a Hivatal teljesítette adatszolgáltatási feladatait, adatszolgáltatását hozzáférhetőbbé tette, stabilizálta szakmai tevékenységét, új fejlesztéseket alapozott meg, valamint elindított

ta a minőségfejlesztési projektet. Elmaradásként értékelte azonban, hogy a szakmai fejlesztési feladatokban egyes helyeken „torlódás” van, a társadalomstatisztikán belüli harmonizációban nem alakult ki konszenzus, és a Hivatal nem lépett előre az adatszolgáltatói kapcsolatok fejlesztésében.

Előadását követően a szakmai kollégiumok vezetőinek beszámolóit hangzottak el. A Nemzeti számlákról a távollévő *Pozsonyi Pál* főosztályvezető helyett *Bedekovics István* főosztályvezető-helyettes adott elő. *Lakatos Judit* főosztályvezető „Lakossági adatgyűjtések”, *Kotulics Tamás* főosztályvezető „Vállalkozásstatisztika”, *Probáld Ákos* főosztályvezető „Szolgáltatásstatisztika – Új feladatok indítása, új adatgyűjtési technika”, valamint *Laczka Éva* főosztályvezető „Mezőgazdasági és környezetstatisztika” címmel tartott előadást.

A délutáni szekcióban *Balogh Miklós*, a KSH elnökhelyettese a Hivatal gazdálkodási helyzetét mutatta be, melyben a 2005. évet értékelte, illetve ismertette a 2006. évi előirányzatot, *Kópházi József*, főosztályvezető szolt a 2006. évi informatikai feladatokról, majd *Németh Eszter* főosztályvezető tartott beszámolót „Tájékoztató – Beszámoló a 2005. évről és a 2006. évi tervek” címmel. *Szemes Mária*, a Veszprémi Igazgatóság igazgatója az igazgatóságok munkaprogramját ismertette, ezután *Dobossy Imre*, főosztályvezető-helyettes tartott beszámolót a KSH Iskola 2006. évi képzési programjáról.

A következő nap programját *Papp Zoltán* főosztályvezető, az MPI vezetője nyitotta meg „Stratégiai fejlesztések mérlege 2005. Hangsúlyok és irányok 2006-ban” című előadásával. Beszámolója után *Kópházi József* „A statisztikai

regiszterek minőségfejlesztése”, Györki Ildikó főosztályvezető-helyettes „Az input folyamatokat támogató alarendszerek fejlesztései”, valamint Szép Katalin főosztályvezető „A statisztikai termékek és munkafolyamatok minőségbiztosítása” címmel tartottak előadást. Baracza Lajosné főosztályvezető-helyettes a KSH metainformációs rendszerének fejlesztését, Hársfai Ferencné főosztályvezető az új gazdálkodási rendszert, Kárpáti József főosztályvezető-helyettes a KSH VIR rendszerének 2006-os fejlesztési terveit, valamint Bodovics Tamás főosztályvezető-helyettes az infrastrukturális fejlesztéseket ismertette. Az éves munkaértekezletet Pukli Péter, a KSH elnökének „Évértékelés, célok és elvárások 2006-ban” című előadása zárta. Beszámolójában többek között bemutatta a Hivatal 2005. évi humán erőforrás-helyzetét, belső szerkezetét, a KSH átalakításokban elért eredményeit, valamint éves, nagy beruházásait. Kiemelte, hogy a Hivatal sikeresen követte, illetve 2006-ban is követni fogja a „KSH – stratégia, 2005–2008” c. dokumentumban megfogalmazott célkitűzéseit.

Változások az OST-ben. A miniszterelnök 8/2006. (I. 19.) ME határozata szerint a statisztikáról szóló 1993. évi XLVI. törvény 7. §-ának (4) bekezdése alapján – a Központi Statisztikai Hivatal elnöke előterjesztésére – Erdősi Sándort, a Belügyminisztérium Rendészeti Statisztikai Önálló Osztálya vezetőjét; dr. Bordás Istvánt, az Egészségügyi Informatikai Szakmai Kollégium elnökét; dr. Székely Juditot, a Foglalkoztatáspolitikai és Munkaügyi Minisztérium helyettes államtitkárát; dr. Szöllősi Endrét, a Földművelésügyi és Vidékfejlesztési Minisztérium főosztályvezetőjét; dr. Tóth Magdolnát, a Gazdasági és Közlekedési Minisztérium főosztályvezetőjét; dr. Belényi Andreát, a Gazdasági versenyhivatal főtitkárát; Sulyok Jánost, a Honvédelmi Minisztérium Központi Pénzügyi és Számviteli Hivatala főigazgató-helyettesét; dr. Ulicska Lászlót, az Ifjúsági, Családügyi, Szociá-

lis és Esélyegyenlőségi Minisztérium főosztályvezető-helyettesét; dr. Kovácsy Zsombort, az Igazságügyi Minisztérium főosztályvezetőjét; dr. Petőfi Lászlót, az Informatikai és Hírközlési Minisztérium helyettes államtitkárát; Bacskai Józsefet, a Külügyminisztérium főosztályvezetőjét; Bozó Pált, a Környezetvédelmi és Vízügyi Minisztérium főosztályvezetőjét; dr. Stauber Józsefet, a Legfőbb Ügyészség főosztályvezető ügyészét; dr. Barát Máriát, a Miniszterelnöki Hivatal főosztályvezetőjét; Gódnorné Kaló Editet, a Magyar Nemzeti Bank főosztályvezetőjét; Légrády Andrást, a Nemzeti Kulturális Örökség Minisztériuma főosztályvezetőjét; dr. Szemán Felicitást, az Országos Igazságszolgáltatási Tanács Hivatala hivatalvezető-helyettesét; Kozma Lukács Juditot, az Oktatási Minisztérium osztályvezetőjét; Tabák Péter, a Pénzügyminisztérium főosztályvezetőjét; dr. Szépvölgyi Editet, az Országos Egészségbiztosítási Pénztár osztályvezetőjét; Paszternák Józsefnét, a Magyar Agrárkamara főtitkár-helyettesét; Kompaktor Emiliát, a Magyar Kereskedelmi és Iparkamara közgazdasági és képzési igazgatóját; dr. Pásztor Miklóst, a Munkástanácsok Országos Szövetsége Szakértői Irodájának vezetőjét; Tóth Lászlót, az Általános Fogvasztási Szövetkezetek Országos Szövetsége tanácsosát; Zilahi-Sebess Gézá, a Stratégiai és Közszolgáltató Társaságok Országos Szövetsége tagját; Geiger Ferencet, Budapest XXIII. kerületének polgármesterét; Jávorszky Imrénét, Berzék polgármesterét; Nagy Sándort Újfehértó polgármesterét; dr. Szép Katalint, a Központi Statisztikai Hivatal főosztályvezetőjét; Sándorné dr. Kriszt Évát, a Budapesti Gazdasági Főiskola tanszékvezetőjét; dr. Szilágyi Györgyöt, a Központi Statisztikai Hivatal ny. főosztályvezető-helyettesét; dr. Vita Lászlót, a Budapesti Corvinus Egyetem tanszékvezető egyetemi tanárát; dr. Hajdu Ottót, a Budapesti Műszaki és Gazdaságtudományi Egyetem egyetemi docensét és dr. Rappai Gábor, a Pécsi Tudományegyetem dékánját 2006.

január 1-jétől 2008. december 31-ig terjedő időtartamra az Országos Statisztikai Tanács tagsági teendőinek ellátásával bízta meg.

Indul a NACE-projekt. A Központi Statisztikai Hivatal 2006-ban útjára indítja a „NACE áttérés” elnevezésű projektjét, amelynek vezetője *Schindele Miklós*, a KSH főosztályvezető-helyettese. A projekt feladata a NACE és a TEÁOR változásával összefüggésben új koncepció kialakítása, majd részletes terv készítése a nomenklátúra-váltás statisztikai kezeléséről. A projekt munkáját két alprojekt támogatja, melyek a feladat regisztert érintő, illetve tájékoztatási oldalával foglalkoznak. A munkavégzés során a projektvezető az érintett főosztályokkal tartja a kapcsolatot. A projekt szakmai felügyeletét a Vállalkozás-statisztikai szakmai kollégium látja el.

MST-ülés. A Magyar Statisztikai Társaság Nemzetközi Statisztikai Szakosztálya 2006. február 22-én, a KSH Árvay János-termében tartotta szakmai ülését, melyen *Kőszeginé dr. Kalas Mária* a közgazdaság-tudomány kandidátusa „Finnország és Magyarország fejlődésének összehasonlító statisztikai elemzése” című előadása hangzott el.

Kitüntetés. A Magyar Köztársaság oktatási minisztere a 2006. évi Magyar Kultúra napja alkalmából a felsőfokú oktatás területén végzett magas színvonalú munkájáért, valamint a pedagógiai gyakorlatot segítő kiemelkedő tudományos tevékenységéért *dr. Vita Lászlónak*, a közgazdaság-tudomány kandidátusának, a Budapesti Corvinus Egyetem tanszékvezető egyetemi tanárának Apáczai Csere János-díjat adományozott.

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

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

Lahore, Pakisztán. 2006. április 3–5.

Első nemzetközi konferencia a statisztikáról és az operációkutatásról. (First International Conference on Statistics and Operations Research.)

Információ: Punjabi Egyetem Statisztikai Intézete, *Sohail Chand*

Telefon: +91-42-9231271

Fax: +92-42-9230634

E-mail: coordinator.icsor@yahoo.com,
coordinator.icsor@gmail.com

Ghent, Belgium. 2006. április 19–21.

Második nemzetközi találkozó: a szájegészségügyi kutatás módszertani kérdései: az adatminőség becslése és fejlesztése. (Second

International Meeting Methodological Issues in Oral Health Research: Assessing and Improving Data Quality.)

Honlap: <http://med.kuleuven.be/biostat/conferences/Dental2006/>

Cardiff, Nagy Britannia. 2006. április 24–26.

Az IASS (International Association for Survey Statisticians–Mintavételi Statisztikusok Nemzetközi Egyesülete) Európai Régió Szekciójának Q2006-os konferenciája. (The Q2006 Conference of the European Regional Section of the IASS.)

Honlap: www.statistics.gov.uk/events/q2006

Könyvszemle

David Freedman – Robert Pisoni –
Roger Purves:

Statisztika

TYPOTEX. 2005. Budapest. 809 old.

Az elmúlt évek egyik nagy publicitást kapott statisztikai könyvének magyar fordítása a *Statisztikai módszerek a társadalomkutatásban* alcímmel jelent meg, az ELTE Társadalomtudományi Karának támogatásával, és ebből arra lehet következtetni, hogy a könyv elsősorban szociológus hallgatók számára foglal össze bevezető statisztikai ismereteket.

A *Jerzy Neyman* emlékének ajánlott könyv mondanivalóját 8 részben és 29 fejezetben fejti ki. Az I. rész bevezető jellegű: a kísérletek tervezését, az ott elkövethető néhány hibát mutatja be példák segítségével, egyelőre nélkülözve bármiféle statisztikai eszköztárat. A II. rész a leíró statisztikába vezet be: fő témakörei a hisztogram, az átlag és a szórás, a normális közelítés, a mérési hiba, és némi meglepetésre a pontok és egyenesek ábrázolása, mely utóbbi témakör elemi koordináta-geometriai bevezetőt takar. A III. rész témája a korreláció- és regressziószámítás, míg a IV. rész a valószínűség-számítási alapokat tárgyalja az alapfogalmaktól az egyszerű összefüggéseken és műveleteken keresztül a binomiális formuláig. Az V. rész témája a véletlen ingadozás, melyen belül szól a nagyszámok törvényéről, bevezeti a várható érték és a standard hiba fogalmát, és foglalkozik az elméleti hisztogramok normális közelítésével. A VI., talán a leghangsúlyosabb és legterjedelmesebb rész: először áttekinti a nagy

felmérések néhány módszertani problémáját példákon keresztül, majd foglalkozik a véletlen hibákkal, a százalékarányok pontossága kapcsán a konfidenciaintervallumok készítésének egyszerűbb eseteivel és értelmezésükkel, a foglalkoztatottság és a munkanélküliség felvételeinek alapján a minták torzításával, a nemválaszolással és az átsúlyozás egyes kérdéseivel, valamint az adatminőséggel. Ugyancsak ez a rész foglalkozik az átlag mintából történő becslésével. A VII. rész címe: Valószínűségi modellek, tartalmát pedig két mintaként választott egyszerű modell: a Gauss-féle mérési hiba modell és a Mendel-féle öröklélmélet modellje alkotja. Végül az utolsó, VIII. rész a szignifikanciapróbákat tekinti át. Az alapfogalmak bevezetése után bemutatja az egyszerű t -próbát, majd sort kerít a z -próbara, a legegyszerűbb kétmintás próbákra, valamint külön fejezetben az illeszkedésvizsgálatra, illetőleg a függetlenségvizsgálatra alkalmazott χ^2 próbákra. Az utolsó fejezet hasznos gyakorlati tanácsokkal látja el az itt bemutatott próbákat alkalmazókat. A Függelék a jegyzeteket, a feladatmegoldásokat, a fontosabb statisztikai táblázatokat, valamint a név- és tárgymutatót tartalmazza.

Ha a könyvet értékelni szeretném, talán azazal kell kezdeni, ami nem tetszik benne (megjegyzem, ez a többség). Mindenek előtt nem értek egyet azzal a szemléletével, hogy statisztikát akar tanítani úgy, hogy kínosan kerüli a matematikai alapokat. Hibásnak és hamisnak tartom azt a szemléletet, amelyik szembe kívánja állítani a matematikai képleteket a gondolkodással. Ez egyébként a magyar szakirodalomban is megjelenik (például *Moksony F.* [1999]: *Gondolatok és adatok*. Osiris. Budapest), de ahány-

szor találkozom vele, minden alkalommal le-
szögezem azt, hogy a képletek, a matematikai
modellek is gondolkodás, mégpedig meglehető-
sen absztrakt gondolkodás eredményei. Ennek a
szemléletnek mindössze annyi lehet a racionális
alapja, hogy bizonyára vannak egyesek, akik a
képletek mögé rejtőzve, azokkal dobálódzva és
többnyire felületesen, azokat szakszerűtlenül
használva (s ami ennél talán még rosszabb,
hogy ha a képleteket a számítógéppel helyette-
sítjük) próbálják megspórolni a gondolkodást.
Ez a magatartás természetesen hibás, visszatet-
sző és nyilván kemény bírálatot érdemel, de úgy
gondolom, ha csak ez lebeg a szemünk előtt, és
ezt elkerülendő számúzzuk a formulákat, a gye-
reket is kiöntjük a fürdővízzel együtt. A mate-
matika segítséget, mankót ad, korábbi fázisok
gondolkodását (másokét vagy sajátunkat) sűríti
össze, és lehetőséget teremt arra, hogy ne kell-
jen újra és újra végiggondolni mindazt, amit
egyszer már végiggondoltunk. A matematika
nem a gondolkodás ellensége, hanem annak
igen hathatós támogatója.

A könyv azt sugallja, hogy nem érdemes,
nem kell venni a fáradságot arra, hogy a ma-
tematikát (feleslegesen) megtanuljuk, a nélkül
is lehet statisztikát készíteni. Ez a könyv kísér-
letet tesz arra, hogy megmutassa hogyan. Vé-
leményem szerint pontatlanul, gyakran értel-
metlenül, és mindezt csak egy bizonyos szin-
tig. A tárgyalás ugyanis általában módfelett
pontatlan és pongyola. A fogalmakat nem is
próbálja meg pontosan definiálni (még a ki-
emelt, keretbe foglalt mondatokat se), az állí-
tások gyakran nem helytállóak, hibásak, félre-
érthetők, ritkán utalnak arra, hogy az állítás
milyen feltételek mellett érvényes. Kiváltképp
nehezményezem, hogy a könyv pontatlanul és
hanyagul bánik a mértékegységekkel. Álljon
itt néhány elrettentő példa.

„A statisztika annak mestersége, hogyan le-
het bizonyos rejtelmes kérdésekkel kapcsolato-

san számszerű következtetésekre ... jutni.” (15.
old.) Ilyen rejtelmes kérdés például a mottó sze-
rint az, hogy milyen dalt énekelnek a szirének.

„...elmagyarázzuk, hogyan olvasandó a
hisztogram. Mindenekelőtt, nincs függőleges
tengely...” (52. old.) Ez önmagáért beszél.

„A szórás megmutatja, milyen messze es-
nek egy lista számai az átlaguktól.” (88. oldal)
Pontosabban: átlagosan milyen messze esnek...

„Egy lista számainak durván 68%-a... az
átlagtól egy szórásnyin belül esik...” (88. oldal)
Ez bizony igen durva általánosítása a normális
eloszlással való közelítésnek. Szó sincs arról,
hogy milyen listákra igaz ez, és milyenekre
nem.

„... a korrelációs együttható kiszámítási el-
járása. Számítsuk át mindkét változót standard
egységbe. A korrelációs együttható az így kép-
zett szorzatok átlaga.” (159. old.) Itt csak az ma-
radt le, hogy hogyan képezzük a szorzatokat. Ha
receptkönyvről van szó, akkor az ilyen pontat-
lanságok ehetetlen ételt eredményeznek.

„Ezzel készen vagyunk a standard hiba
becslésére szolgáló ún. „bootstrap” eljárás vég-
rehajtásával.” (421. old.) Az itt bemutatott eljá-
rásnak semmi köze a valódi bootstraphez. Itt
semmi más nem történik, mint a hiba becslése-
kor magából a mintából, az abból becsült érté-
kekből indul ki. Az igazi bootstrap egy számí-
tógépes ismétléses (szimulációs) eljárás.

„A [konfidencia] intervallum alsó és felső
határt ad meg a paraméterre, valamint annak
megbízhatósági szintjét, hogy a valódi érték be-
leesik az intervallumba.” (429. old.) A valódi
érték nem esik sehova, az rögzített, bár ismeret-
len. Ezért ez az interpretáció egyszerűen hibás!

„A nullhipotézis azt az elgondolást fejezi
ki, hogy a megfigyelt eltérést a véletlen okoz-
za.” (529. old.) A nullhipotézis ennél keveseb-
bet mond; arra nézve, hogy mi okozza az elté-
rést, éppenséggel semmit.

„Ha volna igazság a világban, a P -érték]
azt mondaná meg, hogy milyen valószínűség-

gel igaz a nullhipotézis...” (532. old) Mi az igazság? Ez bizony nagy és rejtelmes kérdés, olyannyira, hogy többnyire statisztikával sem válaszolható meg.

Ennek a szemléletnek szükségszerű velejárója az értelmetlenség: az, hogy bizonyos állításokat, amiket viszonylag egyszerűen lehetne matematikai eszközökkel bizonyítani, meg sem próbál magyarázni (mert intuitív alapon nem is igen lehet), hanem megelégszik azzal, hogy „...higgyük el, hogy ez az állítás bebizonyítható, de ne kérdezzük, hogy miért”. Az ilyen állítások tönkreteszik azt a kétségtelen eredményt, hogy néhány fogalom, állítás intuitíve valóban szépen és hihetően magyarázható. Ha az olvasó gyakran találkozik azzal, hogy nem kell rákérdezni a miértre, az egész tanulási folyamat hiteltelenné válhat, és ami marad az egy egyszerű receptkönyv. Álljon itt két ilyen, jellemző állítás:

„De miért pont r a megfelelő szorzótényező? ... A köztes r értékek esetén bonyolultabb matematikai bizonyításra van szükség; de higgyük el, hogy valóban r a használandó szorzótényező.” (191. old.)

„Nem az a kérdés, hogy mit jelent, hanem, hogy hogyan használják” (580. old. mottó.)

Az elmondottak mellett az ilyen tárgyalásmód erősen korlátozott érvényű, mert megfelelő eszköztár nélkül nem lehet messzire eljutni, hiszen a bonyolultabb modellek (amelyek szerencsére napjainkban a társadalomtudományokban is egyre inkább teret nyernek) verbálisan már egyáltalán nem kezelhetők. (Elegendő a demográfiai modellekre, a többváltozós technikákra, vagy az összetettebb mintavételi problémákra utalni.) Ráadásul az igazi baj az, hogy az itt megismert eszközök nem is teremtenek valódi alapot ezek megismeréséhez, hiszen ehhez a statisztika matema-

tikailag megalapozott elméletével kell kezdeni, azaz mindent az elejéről kell újra megtanulni. Ezért ez a szemlélet mindazok számára, akik egy kicsit is értő módon akarnak a statisztikával foglalkozni, zsákutcának tűnik.

Nem tetszik továbbá az, hogy amerikai könyvet fordítottak le magyar hallgatók részére. Egyrészt azért, mert közismert, hogy az ottani kezdő egyetemi hallgatóság többnyire alacsonyabb felkészültségű (még ma is), mint a hasonló korú magyar fiatalok. Ez eleve rányomja bélyegét a színvonalra (szerintem egy átlagos magyar kezdő egyetemistának esetenként nehéz leszállni ilyen alacsony szintre), másrészt olyan témákat is tárgyal (például a koordináta-geometria legegyszerűbb feladatai), amiket már a középiskola első két évében mindenki ennél színvonalasabban megismert. Nem tartom szerencsésnek ilyen alapozó szinten amerikai könyv fordítását azért sem, mert az ottani gondolkodásmód más, mint az európai, mások a problémák, más a környezet, gyakran mások a lényegi kérdések stb. (Ha már fordításban gondolkozott a döntéshozó, egy jó európai alapkönyv talán szerencsésebb lett volna.) De még akkor is: mások a mértékegységek, ismeretlenek a rövidítések, más jellegűek az adatbázisok, és aligha vitatható, hogy ha a hallgatók valós, és számukra érdekes adatokon gyakorolhatnak, nem csupán a módszereket ismerik meg, de közelebb kerülnek választott szakmájukhoz is. Másként szólva, egy magyar könyv, amely valós magyar adatokat közöl példa gyanánt, nem csak a módszereket, de az aktuális magyar valóságot is megmutathatja. Ez a lehetőség bizony így kimarad. Csak egy furcsa példát hadd említsek a más gondolkodási módra: a rögbilabda alakú pontfelhő egy amerikai hallgatónak nyilván szemléletes, de kérdés, hogy a magyar hallgatók közül hányan láttak közelről igazi rögbilabdát?

Komoly gondjaim vannak a mű szerkezetével, az egyes részek egymásra épülésével is. A

fejzeteszerkezet számomra némiképp ötletszerűnek tűnik. Bár a bevezető fejezet szerepe még talán érthető, nehezen tudom elfogadni, hogy a normális közelítés megelőzze a valószínűség-számítási alapokat, vagy éppen a regressziót. Az egyes fejezeteken belül is ilyen, logikailag nehezen indokolható sorrendiség van. Néhány példa ezekre: előbb van a hisztogram, és csak azután a gyakoriság és gyakorisági eloszlás, előbb tárgyalja (felszínesen) a normális eloszlást és csak később, viszonylag részletesen a binomiális eloszlás néhány jellemzőjét, előbb találkozik az olvasó a lineáris regresszió korrelációból származtatott (nem mondom, hogy levezetett) alakjával, mint a standard formával, előbb beszél a medián néhány tulajdonságáról, és csak később definiálja. Lehet, hogy ezek mögött a furcsa sorrendiségek mögött van valami mélyebb logikai fűzér, de azt nem ismertem fel. A statisztika hagyományos, évtizedek óta jól bevált, egymásra épülő elemein nevelkedett olvasó számára ez bizony nehezen követhető.

Bár azt, hogy milyen témakörökkel foglalkoznak a szerzők, és milyenekkel nem, általában nem illik bírálni, mégis megemlítem, hogy ilyen keretekben és ilyen eszköztárral talán nem kellett volna foglalkozni a páros mintákkal, a kiegyensúlyozott félminták módszerével, vagy viszonylag részletesen a valószínűségi modellekkel, hiszen ezek megfelelő mélységű tárgyalása az itteninél jóval nagyobb eszköztárat igényel. Ugyanakkor hiányolok néhány olyan témát, amelyeket ilyen keretek között is lehetne értelmesen tárgyalni, és a társadalomkutatásban is komoly szerepük lehet. Ezek között említem meg például a további egyszerű leíró mutatókat, a decilisek és a decilis eloszlás bevezetését, az aggregáció egyszerűbb eszközeit, az ordinális skálákra (rangszámokra) épített egyszerűbb elemzéseket, vagy a rétegzett mintavétel bemutatását.

A sajátos szemléletmódból adódóan, nem ritkán a pontatlanságok kiküszöbölése érdeké-

ben, a szerzők igen gyakran voltak kénytelenek olykor igen részletes magyarázatot fűzni a főszöveghez, jegyzetek formájában. Ezek a jegyzetek kiegészítéseket, pontosításokat, vagy formális levezetéseket tartalmaznak, s mivel általában terjedelmesek, szétfeszítenék a lábjegyzetek adta keretet. Ezért ezeket külön szekcióba, a Függelékbe helyezték, így használatuk igen nehézkes. Nem tartom jónak azt sem, hogy a szöveget hol komoly, hol tréfásnak szánt ábrák tarkítják. Úgy vélem, egy egyetemi szintű könyvbe nem valók ezek az eszközök. Hasonlóan ellene vagyok a sok mottónak, kiváltképp azért, mert ezek nagy része nem találó, nem segíti a következő fejezet jobb megértését, sőt bizonyos háttér-információk nélkül az átlagos olvasó számára olykor nem is érthető. (Ennek eklatáns példája az a mottó, amelyet a fordító sem értett meg, és ezért – szerintem korrekt módon – nem is vállalkozott lefordítására.) Csak egy példa a szerintem nem idevaló mottókra (a koordináta-geometriai fejezet elé):

„Kérdés: Mit mondott az egyes a pontnak? Válasz: Pont te hiányoztál.” (134. old.)

A sok bíráló megjegyzés mellett azért természetesen egy sor olyan eleme van a könyvnek, amelyek tetszettek, sőt amelyeket esetenként igen kiválóan tartok. Mindenek előtt tetszik a sok példa. Igaz, egy társadalomtudósoknak ajánlott könyvhöz képest talán túl sok az orvosi-gyógyszerészeti-biológiai példa és feladat, de ezek többnyire nem mennek olyan szakmai mélységekig, hogy ne legyenek érthetőek a kívülálló, más szakmákból érkezett olvasók számára. Nagyon jók, kifejezetten élvezetesek a történeti érdekességgel bíró példák (többek között *Galton*, *Laplace*, *de Moivre*, *Gauss*, *Mendel* stb. eredeti gondolatainak rekonstrukciói), hiszen ezek a statisztikai gondolkodás történetének egy-egy igen érdekes metszetét adják. Ugyancsak érdekesek a napi életből (például bí-

rósági tárgyalásokból, választási statisztikákból) merített példák, valamint az amerikai társadalomkutatási eredmények részletes bemutatása. A példákkal kapcsolatban még két pozitívumra szeretnék rámutatni: egyrészt kevés kivétellel valódi adatokon nyugszanak, és ahol nem, a szerzők nem szégyellik bevallani, hogy fiktív adatokat használtak fel. Másrészt jónak tartom azt, hogy negatív példák is bőségesen szerepelnek, azaz a könyv bemutat sok olyan esetet, ahol a statisztika helytelen alkalmazása téves következtetésekre vezetett. Ezek szerintem didaktikailag talán még fontosabbak, mint a pozitív példák.

Tetszetek bizonyos fejezetek, elsősorban a legelső és a legutolsó. Az első fejezetben az tetszik, hogy szándékosan nem kezd neki ijesztő módszertani fejtegetéseknek, hanem könnyed stílusban, mintegy regényszerűen elmesél néhány esetet, amelyek a későbbiekhez jó kiindulópontot adnak ugyan, de nem rettennek el az olvasót. Az utolsó fejezet a szignifikanciaszint alkalmazásaihoz ad gyakorlati tanácsokat. Itt nagyon jól összefoglalja azokat a buktatókat, amelyek a tapasztalatlan alkalmazóra leselkednek, és implicite bár, de körvonalazza azokat a feltételeket, amelyek között ez az eszköztár korrekten alkalmazható. Bár a mottók túlburjánzását már bíráltam, akad néhány köztük (például a 18. fejezetet bevezető *Lippmann*-idézett), amelyek valóban tömören összefoglalják a lényegét, ezzel nagyban segítik a mondanivaló megértését:

„[A normális közelítésben] mindenki hisz, a kísérletezők azért, mert azt hiszik, hogy matematikai tétel, a matematikusok azért, mert azt hiszik, hogy kísérleti tény.” (350. old.)

Befejezésül elismeréssel kell szólni a fordítók és a szerkesztők munkájáról. A fordítók nem csak mondatonként átvitték a szöveget magyarra, de láthatóan igyekeztek tompítani

azokat a nehézségeket, amelyeket az amerikai könyv eltérő gondolatvilága kapcsán már említettem. Tetszik az, hogy megpróbálták a számunkra idegen mértékegységeket SI-rendszerbe átkonvertálni, ami esetenként nem csekély többszámításokat is igényelt. Ahol a fordítás nem tudta volna teljesen visszaadni az eredeti gondolatot, magyarázó jegyzetekkel egészítették ki a szöveget, és párját ritkító az a már említett korrektség, hogy azt a mottót, ami az amerikai olvasók számára bizonyára érthető volt, de a mi ismereteink nem elégségesek annak megértésére, meg sem próbálták lefordítani. A fordítók általában jól használják a statisztikai nyelvet. Érdekes kísérlet az outlier fordítása *magányos értékként*, bár lehet, hogy ez nem fog elterjedni a szakmai nyelvben. Ugyancsak érdekes, bár talán vitatható az urnamodellek *dobozmodelleként* történő emlegetése, bár megjegyzendő, hogy a fordítók az átmenet kulcsát is megadják. Egyes szakmákban (például ökonometria) a *reziduálisok* helyett inkább reziduumokat, a *megfigyelt szignifikanciaszint* helyett a *p-értékre* empirikus szignifikanciát, *homoszcédasziticitás* helyett, pedig homoszkedasziticitást használnak, de ezek olyan árnyalatok, amelyeket egy fordításnál nem is lehet figyelembe venni. Egyedül talán azt sajnálom, hogy az igen félrevezető *ökológiai korreláció* elnevezést (ami persze így honosodott meg társadalomtudományi körökben) nem próbálták meg valami kevésbé félrevezetőre cserélni. Mindent egybevéve azonban úgy gondolom, hogy a fordítók igen jó és alapos munkát végeztek. Hasonlóképpen dicséret illeti a szerkesztők munkáját, bár egy-két hivatkozott jegyzet eltűnt (652., illetve 657. old.), másoknak pedig a számozása keveredett meg egy kicsit, és a képletekben is akad kisebb sajtóhiba (például 660. old). Mindazonáltal úgy gondolom, hogy az ilyen hibák egy hatalmas méretű műben elkerülhetetlenek, és ezektől eltekintve valóban nagyon gondosan, szép kivitelben jelentették meg ezt a könyvet.

Összességében a korábban részletezettek alapján nyugodtan állíthatom, hogy bár sok elemében vitatható, mégis nagyon figyelemre méltó könyvet tart az Olvasó a kezében. Azt azonban egy pillanatra se felejtjük el, hogy ez a könyv nem statisztika tankönyv, hanem mese

a statisztikáról. Tanulságos mese, amit minél több statisztikusnak érdemes megismerni.

Hunyadi László

a Statisztikai Szemle főszerkesztője
E-mail: laszlo.hunyadi@ksh.hu

Folyóiratszemle

Fairlie, R. W.:

Vállalkozáselemzés az 1979. évi longitudinális felvétel alapján az Egyesült Államokban

(Self-employment, entrepreneurship and the NLSY79.) – *Monthly Labor Review*. 2005. 2. sz. 40–47. old.

Értékes adatsorok nyerhetők a vállalkozásokra vonatkozóan az Egyesült Államokban, a fiatalok körében 1979-ben végzett longitudinális felvételtől (National Longitudinal Survey of Youth – NLSY79). Összesen 12 686 személy adatait tartalmazza az NLSY79 reprezentatív mintája, olyan férfiakét és nőkéit, akik 1979-ben a 14–22 éves korosztályba tartoztak és ez volt az első ilyen összeírásuk. Ezt követően 1979 és 1994 között évente, majd 1996-tól kétfévente követték egymást a felmérések. A vállalkozáselemzés céljára kiemelt minta nem tartalmazza a hadseregben 1978. szeptember 30-án szolgálókat (összesen 1 280 főt).

A szerző ismerteti azokat az adatköröket, amelyek leírják a személyek jellemzőit. Három nagy etnikum határozható el: az afro-amerikaiak, a spanyolajkiak és a fehérek. A felvételekben a szokásos demográfiai kérdések szerepelnek, továbbá a válaszoló gazdasági helyzetét,

családi hátterét, képzettségét, pszichikai állapotát jellemző adatok. Az NLSY79 felmérésorozat sajátos adatokat is tartalmaz, például a válaszoló vagyoni helyzetére, esetleges büntetéseire, a katonai minősítő vizsgálat szerinti besorolásra (Armed Forces Qualification Test – AFQT) vonatkozóan. Felmérik továbbá, hogy a válaszoló milyen módon lépett be a vállalkozásba, illetve vált ki onnan, hány évig volt ilyen jövedelme, véleménye szerint milyen nehézségek gátolják, hogy (ha akarna) vállalkozást folytasson. Az önfoglalkoztatók körében felmérték, illetve becsülték a munkajövedelmet, a munkával való elégedettséget, a működtetett nettó eszközértéket, az eszközök hozamait.

A cikk fogalmi meghatározást ad az önfoglalkoztatókra: akik magukat ilyennek minősítették az alapján, hogy saját vállalkozásként gazdasági tevékenységet, kereső foglalkozást folytatnak, illetve farmot tartanak fenn, és ez a munkavégzési besorolás vonatkozik a jelenlegi, illetve legutóbbi munkapiaci helyzetükre. A nem fizetett segítő családtag eszerint nem sorolható az önfoglalkoztatók közé. Rendszerint meghatározzák azt az éves munkaidőhatárt, amely alatt nem tekinthető a válaszoló önfoglalkoztatónak, például az iskola melletti munkavégzés, ha az a megelőző naptári évben nem érte el a 300 órát.

Megjegyzés. A *Folyóiratszemlét* a Központi Statisztikai Hivatal Könyvtár és Levéltára (*Rettich Béla*) állítja össze.

A cikk táblázatosan közli, hogy a férfiak és a nők az említett etnikai csoportok mintájában milyen arányban tekinthetők önfoglalkoztatónak. A vállalkozási hajlandóság és lehetőség több tényezőtől függ, ezek közé sorolja a szerző a válaszoló nemét, etnikai csoportját, családi állapotát, valamint életkorát. A jól felmérhető jellemzők mellett, közvetve, azt is kimutatták, hogy a vállalkozás mennyiben függ a kockázatok kezelésétől, a személyes adottságoktól, az önállóságnak tulajdonított értéktől, illetve a keresetet biztosító munkavállalással kapcsolatos elkötelezettségtől.

Az NLSY79 felvétel a büntetőügyekre is tartalmaz értékelhető adatokat. A szerző úgy véli, hogy akik „priusszal” jelentkeznek a munkapiacra, azokat megkülönböztetések érnék, ezért vonzóbb számukra az egyéni vállalkozás. Azok a negyvenedik életévükhöz közelítők, akik egykor illegális cselekményre voltak kaphatók (például kábítószer-terjesztés), vagy akiket a bíróság jogerősen elzárással büntetett, azok körében mind nagyobb az önfoglalkoztatás aránya.

A cikk diagramm alapján elemzi a férfiak és a nők önfoglalkoztatási arányának alakulását az 1979 és 2002 közötti időszakban, vagyis a mintába tartozó személyek 22. életévétől a 42. életév felé haladva. A szerző részletesen kifejti, hogy más munkapiaci státusokkal kapcsolatban milyen „átlépési” esetek jellemzők, egyrészt az önfoglalkoztatásra áttérők, másrészt az onnan kilépők arányaival. Az adatsorok arra utalnak, hogy a vállalkozáshoz felhalmozott eszközök nagy szerepet kapnak az egyéni döntésekben, azok kezdhetnek ilyen önálló tevékenységbe, akik lehetőséget látnak befektetéseik finanszírozására, és arra, hogy az elvárt megtérülés elérhető lesz. Igazolható olyan tendencia, hogy azok növelik a vagyonukat, akik vállalkoznak, és kevésbé helytálló, hogy akinek kellően sok eszköze halmozódott fel, az valószínűleg hajlamosabb vállalkozásba kezdeni.

Az adatsorok alapján elemezhetők az önfoglalkoztatás jövedelmi kilátásai. Módszertani kérdés, hogy a felmérést megelőző évben szerzett bevételek alapján milyen jövedelmeket összesíthetők. A cikk táblázatosan hasonlítja össze a munkaviszonyból származó, valamint az önfoglalkoztatással elért éves jövedelmek átlagos értékeit a férfiak és a nők mintáiban. A szerző külön is vizsgálja a befektetett eszközök megtérülését, a vállalkozás, illetve a farm éves működési eredményei alapján. Értelmezési gond az, hogy az önfoglalkoztatók túlnyomó része inkább bérként használja fel a jövedelmeit, nem jellemző, hogy elhatárolnák az üzleti befektetéseik megtérülését.

A felmérés említett rétegei szerint is elvégezték a jövedelem alakulásának elemzését. A mintába tartozó férfiak, akik hátrányos helyzetű családdal jellemezhetők, önfoglalkoztatóként nagyobb keresetet értek el, mint a mintába tartozó munkavállalók, viszont a női minta esetén éppen fordított a helyzet, a munkavállalók átlagos keresete a nagyobb. A férfiak körében, a 22. életévtől kezdve, mindenkor nagyobbak az önfoglalkoztatók átlagos éves jövedelmei, mint a munkavállalóké. A mintába tartozó nők esetén a 30. életév alatt a munkavállalók éves keresete a nagyobb, de az önfoglalkoztatók gyorsabban képesek növelni a jövedelmeiket és az említett kor fölött már többet keresnek. A válaszolók etnikai jellemzői szerint is megállapíthatók ezek a szintkülönbségek mindkét nem esetén, bár az életkori határok esetenként változnak. Az önfoglalkoztatók veszteségeinek kockázatát is figyelembe kell venni, különösen a kezdeti években, amikor még nincs elegendő vállalkozási tapasztalatuk és sok kezdő feladja vállalkozását.

Modellszámításokat végeztek a képzési ráfordítások megtérülési viszonyaira, és ebben az önfoglalkoztatók kedvezőbb helyzetben vannak, mint a mintába tartozó munkavállalók. Mélyebb elemzés is készült, amely figyelembe veszi a

szülők iskolai végzettségének, az olvasottság stb. hatását is.

Akik az 1979-es felmérésben a 14-22 éves korcsoportokba tartoztak, azok 2002-ben már 37-46 évesek lettek, így kellően sok adattal jellemezhető a vállalkozást érő fontosabb gazdasági, társadalmi hatások. Megfigyelhető például a válaszolók munkával való elégedettsége, felhalmozott eszközeik nettó értéke, a munkaerőpiaci esélyeik alakulása. A szerző az azonos korúak csoportjával összehasonlítva vizsgálta azokat a 22. és 26. életévük közötti önfoglalkoztatottakat, akik 27. életévüktől munkavállalóként érvényesültek. Akiknek volt vállalkozási tapasztalatuk, azok gyorsabb ütemben jutottak több keresethez, mint azok, akik korábban is munkavállalók voltak. Tény viszont, hogy a munkaadók rendszerint más követelményeket szabnak, mint amelyek korábban a vállalkozáshoz kapcsolódtak, ezért bizonyos begyakorlási időre is szükség van, az alkalmazás feltételeinek mélyebb megismeréséig.

Nádudvari Zoltán

a Központi Statisztikai Hivatal főtanácsosa
E-mail: zoltan.nadudvari@ksh.hu

Bellmann, L. – Hilpert, M. – Kistler, E. – Wahse, J.:

A demográfiai átalakulás hatása a munkaerőpiacra és a vállalatokra

(Herausforderungen des demografischen Wandels für den Arbeitsmarkt und die Betriebe.) – *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 2003. 2. sz. 133–147. old.

Németországban a demográfiai átalakulás egyaránt hat a munkaerőpiacra és a vállalkozásokra. A munkaerőpiacot érintő demográfiai tényező, hogy a népesség elkerülhetetlen csökkenése és öregedése viszonylag hamar

munkaerőhiányhoz vezet. A korösszetétel eltolódása az idősebb korosztály felé a vállalkozások életében gazdasági növekedést, az innovációs képesség mérséklődését eredményezi.

A szerzők a cikk első részében a demográfiai változásokat, ezeknek a munkaerőkínálatra gyakorolt, nem mindig lineáris hatásait elemzik, a második részben pedig az IAB (Institut für Arbeitsmarkt- und Berufsforschung) vállalati panelfelvétel alapján a keresleti oldal jellemzőit, az idősebb munkavállalókkal kapcsolatos vállalati véleményeket és magatartást tekintik át.

A munkaerő-piaci fordulatról korábban elterjedt nézetek általában a népesség-előrejelzések valamely szélsőséges változatára építenek. Amennyiben egy átlagos, hihető népességalakulási forgatókönyvet veszünk figyelembe, akkor nem reális feltételeznünk, hogy 2020 előtt a munkaerőpiac beszűkül, és hosszabb távon sem a munkaerő hiánya, hanem inkább az elöregedése jelenti a legnagyobb problémát.

Az ok, amiért a politika napirenden tartja ezt a kérdést, hogy az OECD és az Európai Bizottság is felszólította Németországot, tegyen meg mindent az idős munkavállalók foglalkoztatásának növelésére, ezáltal a demográfiai folyamatok kezelésére és a gazdasági növekedés biztosítására, mert az 55 évesnél idősebbek foglalkoztatási aránya alacsony.

Ez a megállapítás azonban a szerzők szerint csak féligazság. Az idősebb korosztályok foglalkoztatási aránya a munkaerő-felmérés adatai szerint Németországban – a volt keleti tartományokkal együtt is – az 1995 és 2000 közötti öt éves periódusban az EU-15 átlagánál kismértékben magasabb volt. Ugyanakkor kedvezőtlen jelenség, hogy az időszak alatt nem nőtt az említett korosztályt tekintve a foglalkoztatási arány, hanem kismértékben visszaesett. Egyes elemzők úgy vélik, hogy a németek túl korán vonulnak nyugdíjba, a nyugdíjazottak életkora átlagosan 60,2 év, a csökkent

munkaképesség miatt nyugdíjazottakat ide nem számítva pedig 62,2 év. Ez azonban annak köszönhető, hogy a munkaképesség idő előtti csökkenése részben a magas és kimerítő munkaintenzitásból adódik, részben nemcsak korfüggő, hanem bizonyos szakmákban – egyes képességek elhasználódása miatt – nagyobb arányban van jelen.

Figyelembe kell venni, hogy a foglalkoztatottakra háruló ún. eltartási teher ezekben az években magas. A 65 év feletti népesség aránya a munkavállalási korúakhoz viszonyítva valóban magas és növekvő, ugyanakkor a gyermekkorúak aránya nagyon alacsony, így összességében az aktív népességre háruló eltartási teher kisebb, mint amekkora például az 1970 és 1975 közötti időszakban volt.

A tévhitek egyik leghangsúlyosabb eleme, hogy a demográfiai folyamatok következtében már a 2010 utáni években munkaerőhiány állhat elő Németországban. Tény, hogy a lakosság öregedő, a teljes belföldi népesség fogyását csak a bevándorlási többlet ellensúlyozza. A népesség korfájának elemzése azonban azt mutatja, hogy a munkavállalási korú, azaz a 20–65 éves népesség együttes száma 2020-ig, amíg a II. világháború után született nagy létszámú korosztályok a munkaerőpiacon vannak, nem csökken. Az előreszámítás szerint – tízéves periódusokat tekintve – 2050-ben esik vissza jelentősen a munkaképes korú népesség száma. A munkaerőpiac egyensúlyának vizsgálatához azonban a népességszámon kívül egyéb tényezőket is célszerű figyelembe venni. Így például a nyilvántartott és rejtett munkanélküliség csökkentésének szándékát, a munkanap hosszát, melynek csökkentése időről-időre újból felmerül stb.

A következő fejezetben azt vizsgálják a szerzők, hogy az eddigi folyamatokból, intézkedésekből mi használható a jövőbeli problémák kezelésénél. Leszögezik, hogy a foglalkoztatottak idősödése nem új, hirtelen felmerü-

lő jelenség – hiszen a II. világháború előtt született, viszonylag népes korosztályok hasonló helyzetet okoztak –, amelyet eseti módon kezelni kell, hanem olyan állandó folyamat, amellyel hosszú távon számolni kell.

Az elmúlt évtizedek tendenciáiból látható, hogy a tényleges nyugdíjba vonulási életkor már 15 éve nem csökken tovább, ezzel szemben a csökkent munkaképesség miatt egyre korábbi életkorban esnek ki a munkaerőpiacról a dolgozók. Ez az egészségügyi ellátás, betegség-felismerés problémáira világít rá. Sokat nőtt a munkanélkülivé válás miatt nyugdíjba vonulók aránya, amely a vállalatok struktúraváltásából is adódik.

Ma tényként kezelik, hogy a korai nyugdíjazás a munkavállalónak is érdekében áll. Ez az esetek egy részében valóban igaz, azonban felmérések szerint a nyugdíjazottak nem kis része még tovább dolgozna, ha erre lenne lehetőség.

A munkanélküli-segély könnyített feltételei szinte felhívást jelentenek a munkaadóknak, hogy az idősebb munkavállalókkal megállapodva, az ő elküldésükkel hajtsák végre a szükségesnek vélt fiatalítást, generációváltást. 2002-ben mintegy 350 ezren váltak ki ilyen módon a munka világából.

Megállapítható, hogy az eddigi gondolkodással, eszközökkel folyamatosan, hosszú távon nem lehet kezelni az idősödő munkavállalók problémáját, már csak pénzügyi okokból sem.

A munkaadók személyzeti politikájának eszköztárában jelen van az időskorúak kiszorítása a munkaerőpiacról, ennek vélt alapja az idősebbek csökkenő teljesítőképessége, termelékenység. A teljesítőképesség ma a legfontosabb szempont a vállalati döntéseknél, a felvételnél, továbbfoglalkoztatásnál, felmondásnál. Ez az érv a politikai vitákban is gyakran hallható, amikor korai, vagy előnyugdíjazásról, nyugdíjkorhatárról esik szó. Végső soron az idősebb munkavállalók sorsa nagymértékben

függ attól, hogy a vállalati személyzeti politika miként ítéli meg a teljesítménycsökkenést.

A gerontológiai kutatások már rég szembe fordultak azzal az előítéllettel, hogy a kor és a teljesítőképesség között determinisztikus összefüggés állna fenn. Korrektebb, ha a teljesítőképesség egyes tényezőit vizsgáljuk. Az IAB vállalati panelfelmérés első alkalommal 2002-ben vizsgálta reprezentatív mintán, különbözőnek-e a fiatal és idősebb munkavállalók a teljesítőképesség egyes tényezői szerint. A fontosnak tartott tényezőket a megkérdezettek válasza alapján állították össze.

A legfontosabbnak a vállalatok képviselői a munkamorált és a minőség tudatosságot jelezték, majd a rugalmasságot, a munkatapasztalatot és lojalitást. Csak ezek után következett a tanulási képesség, a csoportszellem, a pszichikai terhelhetőség. A skála végén foglaltak helyet az elméleti képzettség, a fizikai terhelhetőség és a kreativitás.

A vállalatok véleménye szerint mindössze négy tényezőben különböznek jelentősen a fiatal és az idős munkavállalók: a fiatalok vannak előnyben a fizikai teljesítőképesség, a tanulási képesség, az idősebbek a munkamorál, és a minőségi tudatosság terén. A munkatapasztalat természetesen az idősebbeknél nagyobb mértékben van jelen. A többi tényezőben a megkérdezettek nem láttak különbséget a fiatal és idős munkavállalók között.

A válaszok súlyozott összegzésének eredménye, hogy az idősebb munkavállalók teljesítőképessége semmivel sem rosszabb a fiatalabbakénál. Általános következtetés, hogy a legeredményesebbek az életkor szerint vegyes összetételű munkahelyi közösségek, amelyek a teljesítőképesség valamennyi fontos jellemzőjével együttesen rendelkeznek.

Waffenschmidt Jánosné

a Központi Statisztikai Hivatal főosztályvezetője
E-mail: janosne.waffenschmidt@ksh.hu

Banister, J.:

Feldolgozóipari foglalkoztatás Kínában

(Manufacturing employment in China.) –
Monthly Labor Review. 2005. júly 11–27. old.

A kínai feldolgozóipar 2002-ben több mint kétszer annyi embert foglalkoztatott, mint a G7-országok összesen. A kilencvenes évek második felében az állami szektor gyengülése és a magánszektor nagymértékű erősödése következtében nagyarányú elbocsátások történtek, nőtt a munkatermelékenység, összességében a foglalkoztatás csökkent. 2000-től a trend megfordult, a versenyképesség emelkedésével ismét egyre több ember talál munkát a feldolgozóiparban. A jelenség vizsgálatát megnehezíti, hogy bár a kínai statisztikai rendszer, az adatgyűjtés és -feldolgozás jelentős átalakuláson ment át az elmúlt években, de a változások még korántsem fejeződtek be. A központi statisztikai szervezetek sikeresen vették át a fejlett ipari országok módszereit a népszámlálás, a háztartás- és a fogyasztásstatisztika terén, de többek között a foglalkoztatás, a gazdasági növekedés és az infláció reális számbavételét még mindig akadályozzák az 1978 előtt kialakított kategóriák és eljárások. Az ipari foglalkoztatottságról szóló adatokat illetően például, más területekhez hasonlóan, a Nemzeti Statisztikai Hivatal (NSH) a felelős. Ugyanakkor a nagyvárosokról kiterjedt adatgyűjtés folyik a Munkaügyi és Szociális Biztonsági Minisztériumban, az Állami Ipari és Kereskedelmi Igazgatóságban, míg a falvak és a kisebb városok foglalkoztatottságáról szóló adatok gyűjtésével és feldolgozásával a Mezőgazdasági Minisztérium foglalkozik. Jelenleg minden vállalat, gazdasági egység, kisüzem és magánvállalkozó év végén köteles jelentést készíteni előző évi „munkaügyi eseményeiről” és az év végi állapotról. Az

adatokat összegyűjtik és egy „jelentési láncon” keresztül továbbítják a kormánynak. Tartalmukat tekintve arról kell információt adniuk a jelentés készítőinek, hogy az adott gazdasági szervezetnél milyen munkakörökben hányan dolgoztak és mennyi volt a bérük havi és negyedévi bontásban, az év végén és átlagosan. A jelentés elkészítéséhez szükséges instrukciókat a Nemzeti Statisztikai Hivatal 1990-ben kiadott és 1998-ban, majd 2002-ben kiegészített kiadványa tartalmazza.

A legutolsó, rendelkezésre álló 2002-es, az NSH által kibocsátott adatok szerint Kínában 83 millió ember talált munkát a feldolgozóiparban, ezt a számot egészíti ki a Mezőgazdasági Minisztérium adata, mellyel együtt a gazdasági ágazatban foglalkoztatottak teljes száma mintegy 109 millióra tehető. A 83 millióból 38 millió ember nagyvárosokban dolgozott, melyből 30 millióan vállalatoknál voltak alkalmazásban, 8 millióan pedig magánvállalkozók, vagy magántulajdonban levő kisüzemek dolgozói voltak. A 30 millióból 29 millió végzett valójában valamilyen gazdasági tevékenységet, 16 millió a magánszektor, 10 millió az állami szektor, 3 millió pedig a szövetkezeti szektor gazdasági egységeiben. A magánszektor így összességében 24 millió embernek ad munkát. A feldolgozóipari foglalkoztatottság trendje a városokat tekintve a következő képet mutatja: 1978 és 1995 között rendkívüli növekedés mutatható ki (53 millióról 98 millióra), majd ezután 2000-ig 80 millióra csökkent a foglalkoztatottak száma, az ezt követő két évben pedig ismét emelkedés következett. A vidéki feldolgozóiparban (amelybe Kínában a falvak és a kisvárosok tartoznak) kisebb visszaesésektől eltekintve egyenletes volt a foglalkoztatottság növekedése (2002-ben 45 millió fő).

A kínai munkaügyi statisztika kategóriái azonban korántsem egységesek, az adatgyűjtést végző szervezetek különböző elnevezé-

seket és definíciókat használnak, a statisztikai jelzőszámok tartalma pedig az elmúlt 20-25 évben többször változott. 1998-ban például minden foglalkoztatottsági mutató értéke váratlanul nagymértékben csökkent, kivéve a magánszektorét. Ennek oka az volt, hogy ettől az évtől kezdve azokat, akik valamilyen kapcsolatban álltak munkahelyükkel, de tényleges munkát nem végeztek, nem számították a vállalatnál foglalkoztatottak létszámába. Természetesen, ha az 1998 előtti mutatókat is módosítjuk ezzel a csökkentő tényezővel, a foglalkoztatás trendjéről is valóságosabb képet kaphatunk. A kalkulációk elvégzése után azonban még mindig irreálisan nagy az iparban foglalkoztatott munkaerő csökkenése, 1997 és 1998 között például 7 millió fő. Ennek szinte egésze a városi dolgozók köréből került ki. Különösen az állami és szövetkezeti tulajdonú vállalatok dolgozóinak száma mérséklődött, mivel maga a szektor is nagymértékben összezsugorodott, a vállalatokból magánvállalatok és külföldi-kínai vegyes tulajdonú cégek lettek. A vidéki ipari és a városi magánszektorban foglalkoztatott dolgozók száma viszont emelkedett. A munkaerőkínálat rendkívüli emelkedése az ipari termelés hatékonyságának és versenyképességének növekedésével járt együtt, ugyanakkor a nyugati országokhoz hasonlóan sok zűrzavart okozott a családok életében.

Ha összehasonlítjuk a hivatalos statisztika adatait a 2000-es népszámlálás eredményeivel, kiderül, hogy a hivatalos statisztikai adatgyűjtési rendszer sok millió fővel kevesebb foglalkoztatottat mutat ki, mint a népszámlálás. A minisztérium és a statisztikai hivatal adataiból hiányoznak a feketemunkások és más „informálisan” foglalkoztatottak. A vidéki háztartások ugyanakkor sok esetben mezőgazdasági munkaerőnek jelölték meg azt, aki részmunkaidőben, vagy az év egy részében az iparban dolgozott. A tévedést erősítette, hogy a foglal-

koztatottságra vonatkozó kérdés úgy szólt, hogy hol dolgozott a családtag október utolsó hetében, ami a betakarítás csúcsidejének számított. Az eltérés az összes foglalkoztatott esetében mintegy 80 millió, a feldolgozóiparban pedig 8 millió fő volt. A különbségeket az is növelte, hogy a hivatalos statisztikában az egyéb foglalkoztatott kategóriába 56 millió fő került, a népszámlálás során pedig szinte mindenkit besoroltak, valamilyen szokásos gazdasági ágazatba. A 8 milliós feldolgozóipari különbségből mintegy 2-3 milliót megmagyaráz a részmunkaidősök számbavétele. A népszámlálás során, aki egy óránál többet dolgozott az adatfelvételt megelőző héten, az már foglalkoztatottnak számított. A vállalati jelentéseken alapuló állami statisztikában az számít részmunkaidősnek, aki egy évnél rövidebb ideje dolgozik. Ezt az adatot egyébként csak 2002-re ismerjük a városi ténylegesen munkát végzők esetében 0,47 millió fő (a 10 millió 5 százaléka). A kétféle adatforrás közötti különbséget az okozhatja, hogy a népszámlálásnál 15, a vállalati jelentésekben pedig 16 év az alsó korhatár és a felső korhatár érvényesítése is bizonytalan a statisztikai hivatal és a minisztérium adataiban.

Újabb problémát jelent a városi és vidéki foglalkoztatottak kategóriája. A hivatalos statisztikában városi dolgozónak tekintik, akinek nagyvárosokban van a munkahelye, a népszámlálási adatokban pedig inkább a lakóhely számított a besorolásnál. A helyzetet tovább bonyolítja, hogy a statisztikai hivatal sok nagy múltú állami cég foglalkoztatottait akkor is városinak tekinti, ha a cég vidéken helyezkedik el.

A statisztikai módszertan átalakításának legújabb törekvései szerint a földrajzi kategorizálásnál az adminisztratív irányítás helyett a tényleges helyszín játszik szerepet. A munkaügyi statisztika problémája Kínában, hogy a statisztikai hivatal adatai valószínűleg nem veszik figyelembe a mezőgazdasági minisztéri-

um falusi és kisvárosi ipari foglalkoztatottságról szóló adatait. Ez utóbbi egyébként 1978 óta létezik a teljes vidéki iparra vonatkozóan, a feldolgozóipart először a 2002. évben különítették el. Az idősorok értékelését nehezíti, hogy a közölt adatok 1997-re hirtelen csökkenést, majd 1998-ra váratlan emelkedést mutatnak. A „jelenség” magyarázata az lehet, hogy az 1997. évre vonatkozóan a mezőgazdasági minisztérium változtatott a statisztikai kategóriák tartalmán. 1998-ban viszont a statisztikai hivatal rendelte el, hogy bizonyos árbevétel alatti vállalatok nem kötelesek jelentést készíteni, így ezek a hivatalos statisztikából kimaradtak és valószínűleg a vidéki foglalkoztatottak számát növelték.

A mezőgazdasági minisztérium ipari, munkaügyi adatait tekintve nem világos, hogy pontosan milyen körre terjed ki, hogyan kezeli például a részidős foglalkoztatást. Az említetteken kívül még sok más bizonytalanság is terheli a munkaügyi statisztikát. Például a 2000. évre vonatkozóan, a statisztikai hivatal és a munkaügyi minisztérium adatai szerint, a feldolgozóipari foglalkoztatottak száma 8 millióval kevesebb, mint a népszámlálás alapján, ha azonban a vidéki ipari foglalkoztatottakat is hozzáadjuk a hivatalos statisztikához, akkor 20 millióval nagyobb számot kapunk. Erre az évre tehát a hivatalos statisztika áll közelebb a census adataihoz. Ha viszont az 1995-ös évet nézzük, amikor ipari census volt az országban, akkor a vidéki ipari foglalkoztatottság adataival kiegészített városi feldolgozóipari foglalkoztatottság (128 millió fő) majdnem megegyezik a census adatával (131 millió fő). Mindent egybevetve, úgy tűnik, hogy a hivatalos statisztika és a mezőgazdasági minisztérium együttes adatai talán jobban közelítik a valós helyzetet, mint a statisztikai hivatal és a munkaügyi minisztérium adatszolgáltatása. Az agrárminisztériumban külön részleg foglalkozik a vidéki vál-

latokról, közöttük a foglalkoztatottságról szóló adatok gyűjtésével.

Sajátos problémát jelent a négy, főként exportra termelő feldolgozóipari régió dolgozóinak besorolása. Ezek létszám- és béradatait jelenleg a statisztikai hivatal gyűjti és sorolja be főként a vidéki ipari foglalkoztatottak csoportjába. A szerző becslése szerint, ha a mezőgazdasági minisztérium előírásai szerint gyűjtenék ezekben a szektorokban az adatokat, jóval magasabb értékeket kapnának, mint a jelenlegi 24 millió fő.

A hivatalos statisztika furcsa módon Kínában nem választja külön a migráns dolgozókat, akik főleg faluról vándorolnak a nagyvárosokba és az exportzónákba. Számukra csak a népszámlálás adataiból következtethetünk: 2000. november 1-jén 14,6 millióan tartoztak ebbe a kategóriába, ami a városi foglalkoztatottak 25 százalékát jelenti. A munkaerő-vándorlást még a legfejlettebbekhez tartozó shanghai exportzóna statisztikája sem kezeli jelentőségének megfelelően, egy 2003-as vizsgálatból készült kiadvány azt mutatja, hogy a hosszú távú bevándorlás adatai nem kerülnek bele a statisztikai jelentésekbe.

Összefoglalóan elmondhatjuk, hogy Kína gazdaságában a feldolgozóipari foglalkoztatottság fontos szerepet tölt be. A létszám a G7 országokban és Kínában is csökkenő tendenciát

mutat, ami a gazdasági verseny és a termelékenység emelkedésének köszönhető. A feldolgozóipar részesedése az összes foglalkoztatott népességen belül Kínában jóval magasabb, mint a fejlett országokban, ami három okra vezethető vissza: az ország feldolgozóipara még mindig meglehetősen munkaintenzív, az ágazat termékei az egész világon rendkívül versenyképesek és nem csak az alacsony bérek miatt, valamint a hatalmas hazai piac kielégítése is a foglalkoztatottak rendkívül magas létszámát igazolja. A kínai munkaerőstatisztikai-rendszer továbbfejlesztéséhez, az adatok helyes értékeléséhez még számos feladatot kell megoldani, ilyenek például a munkaerő-vándorlás számbavétele, a vidéki feldolgozóipari foglalkoztatottság adatgyűjtési rendszerének tökéletesítése, a különböző adatforrások ellentétes adatainak felderítése, munkaerő-felmérések megszervezése és gondos végrehajtása, a 2005-re tervezett, 2004-re vonatkozó gazdasági census eredményeinek részletes elemzése. A munkaügyi statisztika módszertani kérdéseit egyébként a szerző a jelenlegi cikknél jóval részletesebben tárgyalja a <http://www.bls.gov/fls/chinareport.pdf> webcímen található angol nyelvű tanulmányában.

Dévai Péter

a KSH Könyvtár és Levéltár osztályvezetője
E-mail: peter.devai@ksh.hu

Kiadók ajánlata

RÉNYI A. [2005]: *Ars Mathematica*. Typotex Elektronikus Kiadó. Budapest.

A kiadó egy kötetben jelenteti meg a neves matematikus, a magyar valószínűség-számítási iskola megalapítójának „Dialógusok a matematikáról”, „Levelek a valószínűségről” és „Napló az információelméletről” címen koráb-

ban már megjelent írásait. A szerző a görög hagyományhoz visszatérve ismerteti a matematikai elméleteket. Írásai ma sem vesztek aktualitásukból. Mivel Rényi kiváló népszerűsítő író volt, munkái a téma iránt ma érdeklődő számára is hasznos ismeretekkel szolgálhatnak. A munka címadó írása, amely bevezető is, vall a szerző kutatói és tanári elveiről.

GISBERT, S. (szerk.) [2005]: *Matlab*. Typotex Elektronikus Kiadó. Budapest.

A munka bevezet a numerikus és a statisztikai módszerek használatába. Részletesen bemutatja az egy- és kétdimenziós grafikák elkészítésének módját a Matlab programrendszerben. A módszerrel olyan mérnöki, természettudományos és közgazdasági feladatok oldhatók meg, melyek analitikus megoldása már lehetetlen. A könyvben minden bemutatott Matlab-utasítást számos példa szemléltet, valamint tartalmaz feladatokat és a címszavak jegyzékét is.

BLUMAN, A. [2005]: *Probability densified*. (Sűrített valószínűségek.) McGraw-Hill. New York.

Sok ember nem érti a valószínűségelmélet alapjait, amit a lottózók, lóversenyek és kaszinók növekvő látogatottsága is jelez. De a valószínűségvizsgálat és a statisztika ma társadalomban egyre fontosabbá válik. Mindenkinek szüksége van arra, hogy megértse a valószínűségelméletet, hiszen az általános iskolai oktatásban és a középiskolákban is tanítják alapjait. A valószínűségelmélet a lehetőségek véletlenszerű megvalósulásával foglalkozik. Bár nem lehetséges, hogy egy egyéni esemény bekövetkeztét biztosra jövendöljük, azt azonban igen, hogy a következményeket hosszú távon meghatározzuk és megjósoljuk annak a valószínűségét, hogy adott esemény bekövetkezik. A könyv a valószínűségelmélet alapjait kötetlen formában mutatja be. Alapvető aritmetikai és algebrai tudás birtokában megismerhetjük az elméletet. Az alapokat magas szintű matematikai formulák és bizonyítások nélkül gyakorlati példákkal, mint érme-, kocka- vagy kártyajátékok és más hétköznapi valószínűségi előfordulások felhasználásával mutatja be a szerző.

SEVERINI, T. A. [2005]: *Elements of Distribution Theory*. Cambridge Series in Statistical and Probabilistic Mathematics. (Az eloszláselmélet alapjai.) Cambridge University Press. Cambridge.

Ez a részletes bevezetés az eloszlások elméletébe statisztika elméleti kurzusok valószínűségelméleti részének tankönyveként használható a mester- és PhD-szintű statisztikus-, biostatistikus- és ökonometria-hallgatók számára. A könyv nem használ mértékelméletet, csupán a matematikai analízis és a lineáris algebra ismeretét tételezi fel. A tárgyalt témák az alapeloszlásoktól és a sűrűségfüggvényektől kezdődően a várható értéken, a feltételes eloszlásokon, a karakterisztikus függvényeken, kumulánsokon, az eloszlásbeli konvergencián és a központi határeloszlás tételein át ívelnek a valószínűség-számítás magasabb szintű fogalmaiig, a csoportstruktúra modellekig valamint az integrálok és az ortogonális polinomok aszimptotikus approximációjáig. A függelék a könyvben használt matematikai definíciók és eredmények részletes összefoglalóját tartalmazza.

YOUNG, G. A. – SMITH, R. L. [2005]: *Essentials of Statistical Inference*. Cambridge Series in Statistical and Probabilistic Mathematics. (A statisztikai következtetéselemélet lényege.) Cambridge University Press. Cambridge.

Ez a tankönyv a statisztikai következtetéselemélet bayesi, frekventista és Fisher-féle megközelítésének fogalmait és fontosabb eredményeit mutatja be hangsúlyozva a közöttük levő különbségeket. A könyv elsősorban felsőbbéves matematikus, illetve a matematikához kapcsolódó tudományterületek hallgatóinak íródott. Az alapozó matematikai elméletet és az arra épülő magasabb szintű

anyagokat egyaránt kezeli, bemutatva olyan modern témákat, mint a bayesi számítások, a magasabbrendű likelihoodok, a prediktív következtetések, a bootstrap módszerek és a feltételes következtetések elmélete. Számos részletes példát tartalmaz, melyek bemutatják egyrészt a formális következtetésemélet alkalmazását valós adatokra, másrészt adalékul szolgálnak ezen tudományterület történeti kialakulásának megértéséhez. Jóllehet a könyv

inkább a fogalmak megértésére mintsem a matematikai részletekre koncentrálna, megtartja a matematikai formalizmus megfelelő szintjét. Minden fejezet a kapcsolódó problémák felvetésével zárul. A könyv (egyetemi oktatósi segédletként) megértése alapvető valószínűség-számítási ismereteket feltételez, ugyanakkor a tárgyalt témák és a statisztikai következtetésemélet előzetes ismerete hasznos lehet, de nem elengedhetetlen.

Társfolyóiratok



A FRANCIA STATISZTIKAI ÉS GAZDASÁG-
KUTATÓ INTÉZET FOLYÓIRATA

2005. ÉVI 78. SZÁM

Gourieroux, C. – Jasiak, J.: Nemlineáris innovációk és impulzusválaszok VaR-érzékenységre alkalmazva.

Latruffé, L. – Picard, P.: A természeti katasztrófák biztosítása: választás a megelőzés és a szolidaritás között.

Lefranc, A. – Trannoy, A.: A generációk közötti kereseti mobilitás Franciaországban: mobilabb Franciaország, mint az Egyesült Államok?

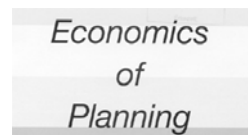
Duguet, E. – Petit, P.: Diszkrimináció a francia pénzügyi szektorban: ökonometriai elemzés a terepi kísérleti adatok alapján.

Bosi, S. – Cazavillan, G. – Magris, F.: A határozatlanság és a komplex dinamika hihetősége.

Joseph, G.: A munkanélküliség allokációs ismérvei és a munkaerőpiac teljesítménye.

Chahuc, P. – Postel-Vinay, F.: Társadalmi státus és a túldolgoztatott fogyasztó.

Pucci, M. – Valentin, J.: A CDD talán kedvező a foglalkoztatásra: vállalatok döntéseinek elemzése környezeti bizonytalanság esetén.



A BIRMINGHAMI EGYETEM FOLYÓIRATA

2004. ÉVI 3–4. SZÁM

Delannay, A. F. – Weill, L.: A kereskedelmi hitel meghatározói a rendszerváltó országokban.

Jevcak, A.: Az állami költségvetés bevételi és kiadási oldala közötti szimmetria szerepe az adóverseny esetleges negatív hatásaira vonatkozó vitában.

Harrison, B. – Paton, D.: Átmenet, a részvénytőke hatékonyságának fejlődése és belépés az EU-ba: Románia esete.

Matousek, R. – Taci, A.: Hatékonyság a bankszakmában: empirikus bizonyítékok Csehországból.

Utkulu, U. – Özdemir, D.: Okoz-e a kereskedelem liberalizálása hosszú távú növekedést Törökországban?

Pavlova, M.- Groot, W.- Merode, G.: Elemzés a minőség, a hozzáférés és az ár jellemzőinek fontosságáról az egészséggondozást igénybevevők számára.

POPULATION

A FRANCIA DEMOGRÁFIAI INTÉZET
FOLYÓIRATA

2005. ÉVI 4. SZÁM.

Verpoorte, M.: A ruandai népirtás: Gikiongoro tartomány esete.

Prioux, F.: A demográfiai helyzet alakulása Franciaországban.

Breton, D. – Prioux, F.: Két vagy három gyermek? A családpolitika és néhány társadalmi-demográfiai tényező hatása.

Peretti-Watel, P. et al.: Hogyan mérjük fel a szeropozitív lakosságot Franciaországban: a VESPA 2003 felvétel.



AZ EGYESÜLT NEMZETEK EURÓPAI
GAZDASÁGI BIZOTTSÁGÁNAK
FOLYÓIRATA

2005. ÉVI 2. SZÁM

Oudhof, K. – Keuzankamp, S.: Változó attitűdök.

Brouwer, R. – Schenau, S. – Veeren, R.: Integrált folyami elszámolás Hollandiában és a vízgazdálkodásra vonatkozó európai irányelvek.

Trott, D. L. – Simpson, A. M.: Számítógéppel támogatott személyi interjú: a bermudai tapasztalat.

Rafalowska, H.: A statisztikai hivatal tekintélyének megszilárdítása hatékony kommunikáción keresztül.

Ostergaard, L.: A statisztikai hivatal tekintélyének folyamatos figyelése.

Zawitz, M. W.: Mi a statisztikai compendiumok jövője a XXI. században?

Shlomo, N. – Waal, T.: Az eredeti információtartalmat megőrző javítások, melyek a statisztikai mikroadatok felfedés elleni védelemben használt perturbációk esetén szükségesek.



AZ EGYESÜLT ÁLLAMOK
MATEMATIKAI STATISZTIKAI INTÉZETÉNEK
FOLYÓIRATA

2005. ÉVI 4. SZÁM

Fan, J.: A pénzügyi ökonometriában alkalmazott nemparaméteres módszerek válogatott áttekintése.

Phillips, P.C. – Yu, J.: Hozzászólás.

Sorensen, M.: Hozzászólás.

Mykland, P.A. – Zhang, L.: Hozzászólás.

Fan, J.: Válasz.

Geyer, C. J. – Meeden, G. D.: Fuzzy és véletlen konfidenciaintervallumok, valamint p -értékek.

Agresti, A. – Gottard, A.: Hozzászólás. Random konfidenciaintervallumok és a közép- p -módszer.

Berger, R. L. – Casella, G.: Hozzászólás. Fuzzy és véletlen konfidenciaintervallumok, valamint p -értékek.

Brown, L. D. – Cai, T. T. – DasGupta, A.: Hozzászólás. Fuzzy és véletlen konfidenciaintervallumok, valamint p -értékek.

Gelman, A.: Hozzászólás. Fuzzy és bayesi p -értékek és u -értékek.

Thompson, E. A.: Hozzászólás. Fuzzy és véletlen konfidenciaintervallumok, valamint p -értékek.

Geyer, C. J. – Meeden, G. D.: Válasz: Fuzzy és véletlen konfidenciaintervallumok, valamint p -értékek.

Jones, B. et al.: Sztochasztikus számítási kísérletek többdimenziós grafikus modellekre.

Aldrich, J.: Fisher és a regresszió.

Barry, D.: Beszélgetés John Hartigan-nel.

Schweizerische Zeitschrift für
**Volkswirtschaft und
Statistik**

A SVÁJCI STATISZTIKAI ÉS KÖZGAZDASÁGI
TÁRSASÁG FOLYÓIRATA

2005. ÉVI 2. SZÁM

Bichsel, R. – Perrez, J.: Kérdéses a bank-kölcsön csatorna: svájci bizonyíték egyedi banki adatok segítségével.

Cattaneo, A. – Winkelmann, R.: Kereseti különbségek a németül és franciául beszélők között Svájcban.

Clausen, J. R. – Meier, C. P.: Követte-e a Bundesbank a Taylor-szabályt? Egy valós idejű adatokon alapuló elemzés.

Dorn, D. – Souza-Poza, A.: A korai nyugdíjazás meghatározói Svájcban.

Reitz, S. – Slopen, U. D.: Makroökonómiai hatások a díjszabásokra: eredmények egy új, nyitott gazdasági makroökonómiai modellből.

2005. ÉVI 3. SZÁM

Asheim, G. B.: Generációk közötti etika erőforrás-korlátok mellett.

Brock, W. – Xepadepadas, A.: Területi elemzés a megújítható erőforrás-kezelés leíró modelljeiben.

Bretschger, L. – Pittel, K.: Innovatív beruházások, természeti erőforrások és generációk

közötti korrektség: jök-e a nyugdíjalapok a fenntartható növekedéshez.

Maradan, D. – Vassiliev, A.: A széndioxid-csökkentés marginális költségei: empirikus bizonyíték országok közötti elemzésből.

Hernlock, M.: Egy differenciális játék az ipari termelés kibocsátásainak alávett természeti tőke kezelésére.

Hwang, U. – Magris, F.: Generációk közötti ellentétek és egy rövid ideig működő kormány erőforrás-politikájának formálása.

Egli, H.: Új módszer a szennyezés modellezésére a környezeti Kuznets-görbe modellekben.

2005. ÉVI 4. SZÁM

Halla, M.: Megélhetés, gondozás és válás.

Jochum, C. – Savioz, M. R.: Egy furcsa állat? A svájci frank árfolyama, mint egy „elfogott” véletlen bolyongás.

Puhani, P. A.: Viszonylagos kínálat és kereslet a szakképzettség területén Svájcban.

Reinowski, E. – Schultz, B. – Wiemers, J.: A továbbképzési programok értékelése egy optimális párosító algoritmus segítségével.

Stutzer, A. – Kienast, L.: Demokratikus részvétel és állami feladat: a női szavazati jog hatása.

Сmamucmuka Statistics

A BOLGÁR STATISZTIKAI HIVATAL
FOLYÓIRATA

2005. ÉVI 2. SZÁM

Tzvetkov, S.: Az olvasóhoz.

Balev, I.: 125 éves a Bolgár Statisztikai Intézet.

Tzvetkov, S.: Statisztika és közgazdaságtan mint egyetemi tárgy – múlt, jelen és jövőbeni fejlesztés.

Petrova, R.: A bolgár statisztika és Bulgária integrálódása az egyesült Európába.

Dimitrov, A.: Hivatalos statisztika és társadalmi-gazdasági kísérletek.

Cholakov, N.: A reneszánsz és a korai statisztikai felfedezések.

Saikova, I.: A liszaboni stratégia Európa gazdasági és társadalmi megújulására és lehetőségek az egyéni innovációs tevékenységekre.

Todorov, V.: Termelési mutatók Bulgáriában amerikai dollárban számítva – fő jellemzők.

Trifonov, T. – Ivanov, I.: Statisztikai elemzés és modellezés wavelet módszerrel.

Tzvetkov, A.: A regionális fejlődés többdimenziós elemzése neurális hálók segítségével.

2005. ÉVI 3. SZÁM

Radilov, D.: A statisztika mint tudomány a XXI. század útkereszteződésénél – jelenlegi kihívások és a jövő.

Kirilov, R.: Statisztikai és információs problémák a hitelműveletek fejlesztésében Bulgáriában.

Tsankova, N.: A statisztikai olvasottság fejlesztésének lehetőségei a Nemzetközi Statisztikai Oktatási Szövetséggel (IASE) együttműködve.

Stoykova-Kunalieva, A.: A statisztikai információ minősége és előállításának ellenőrzése.

Kotzeva, M.: Különböző statisztikai felvételi forrásokból származó adatok kombinálása.

Kosekova, S.: Fogyasztóiár-statisztikák fejlődésének kronológiája.

Grozeva-Koeva, D.: Vállalati körök statisztikai metaadatok iránti igényére vonatkozó felvétel néhány eredménye.

Barbolova, J.: Az európai országok gazdasági fejlődésének mérése és összehasonlítása.

Stoilov, K.: A vállalati körök adatszolgáltatási kötelezettségéből származó teherre vonatkozó felvétel értékelésének főbb eredményei.

Dimitrova, R.: Módszertan kifejlesztése a vállalati szakképzés figyelésére.

2005. ÉVI 4. SZÁM

Kunaliev, T.: A statisztikai információs szolgálat jogi, szervezeti és személyi problémái Bulgáriában.

Ganeva, S.: Stratégia egy integrált statisztikai információs rendszer és információs technológia létrehozására a Bolgár Statisztikai Intézetnél.

Mishev, G.: A statisztika oktatása a közgazdasági felsőoktatásban: döntő tényező a tudósképzésben.

Rangelova, R.: Nemzetközi GDP-összehasonlítások és Bulgária szerepe.

Todorov, T. I.: Zöld GDP: lehetséges módszerek az értékeléshez.

Hristov, A. – Boshnakov, V.: Pénzügyi elszámolások Bulgáriában, a fejlesztés kilátásai.

Kjoseva, N.: A statisztikai adatok minőségének értékelése – módszertani szempontok és megvalósításuk a Bolgár Statisztikai Intézetnél.

Bogdanov, B. – Geron, J. – Toteva, D.: Munkaerő-felvétel: főbb jellemzők és összehasonlítások.



AZ OROSZ ÁLLAMI STATISZTIKAI BIZOTTSÁG FOLYÓIRATA

2005. ÉVI 12. SZÁM

Sokolin, V. I.: Az orosz állami statisztika fejlesztésének kilátásairól.

Darenskikh, Yu. A. et al.: Faktoranalízis egy régió termelésének gazdasági hatékonyságára vonatkozóan.

Smagin, B. I. – Neuymin, S. K.: A lényeg és a módszer a regionális fejlődés mutatóinak meghatározásához.

Bychkova, S. G.: Rendszerszemléletű megközelítés a lakosság életszínvonalbeli regionális különbségeinek statisztikai becsléséhez.

Gerasenko, V. P.: A regionális gazdasági paraméterek szórásanalízise.

Pashintseva, N. I.: A településstatisztikai rendszer formálása.

Chudillin, G. I. – Geniatulina, K. V.: A települések gazdasági fejlődésére vonatkozó többdimenziós becslések szóródása.

Borisova, V. V.: A Rosstat statisztikai regisztere: a fejlesztés irányai.

Tul'chinskiy, B. M.: A statisztikai adatgyűjtés szervezése: ideológia és gyakorlat.

Shemenev, S. A.: A Penza statisztika 170. évfordulója.

Az Orosz Föderáció fő társadalmi és gazdasági mutatói 2001–2005-re.

A 2002-es oroszországi népszámlálás eredményei: életkörülmények.

Sazhin, Yu. V. – Podol'naya, N. N. – Tsyganok, I. S.: A munkaerő reprodukciója a Mordovia Köztársaságban.

Frenkel, A. A.: Az orosz gazdaság fejlődésének előrejelzése 2005–2006-ra.

Wirtschaft und Statistik

A NÉMET SZÖVETSÉGI STATISZTIKAI
HIVATALN FOLYÓIRATA

2005. ÉVI 9. SZÁM

Rath, N. – Braakmann, A.: A GDP 2005 első felében.

Nahm, M. – Philipp, K.: Strukturális adatok a vállalati regiszterből és a vállalati demográfia egyes kérdései.

Luckner, K. H.: Trendek a nagykereskedelemben, 2004.

Walter, K.: Szállítási láncok az intermodális áruszállításban.

Rolland, S.: Megelőzés és rehabilitáció Németországban.

Bechtold, S. – Elbel, G. – Hannappel, H. P.: Az érzékelt infláció mérése Németországban: a vásárlási gyakoriság meghatározása a Szövetségi Statisztikai Hivatalban.

Brachinger, H. W.: Az euró, mint 'teuro'? Az infláció érzékelése Németországban.

2005. ÉVI 10. SZÁM

Sturm, R.: A szövetségi statisztika által a vállalatokra rótt adatszolgáltatói kötelezettség mérése.

Afentakis, A. – Bihler, W.: A 2005-ös évközi mikrocenzus adatainak kiterjesztési eljárása.

Rengers, M.: Az ILO munkaerő-piaci statisztikájának munkanélküliekre vonatkozó havi idősorai.

Ziebach, M.: Építési célú megtakarítások, 2004.

Tümmeler, T.: A vállalatok szolgáltatási igénye.

Hauschild, W. – Schnorr-Backer, S.: E-gészség: az elektronikus egészségügy rendszer Németországban.

Dittrich, S.: Forgalom és adóztatása, 2003.

Rehm, H.: Közpénzügyek 2005. első felében.