

STATISZTIKAI SZEMLE

A KÖZPONTI
STATISZTIKAI HIVATAL
FOLYÓIRATA

SZERKESZTŐBIZOTTSÁG:

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

79. ÉVFOLYAM 8. SZÁM

2001. AUGUSZTUS

E SZÁM SZERZŐI:

Adorján Richárd, a Budapesti Közgazdaságtudományi és Államigazgatási Egyetem PhD-hallgatója; *Gray, Roger*, a Heriot-Watt Egyetem (Edinburgh) tanszékvezető-helyettese; *Hámori Gábor* közgazdász; *Hüttl Antónia*, a KOPINT DATORG Rt. tudományos tanácsadója; *Dr. Kovács Erzsébet* kandidátus, a Budapesti Közgazdaságtudományi és Államigazgatási Egyetem docense; *Nádudvari Zoltán*, a KSH főtanácsosa; *Oravecz Beatrix*, a Budapesti Közgazdaságtudományi és Államigazgatási Egyetem PhD-hallgatója; *Dr. Pozsonyi Pál*, a KSH fősztályvezetője.

*

Bíró Júlia kandidátus, a KSH Könyvtár és Dokumentációs Szolgálat főkönyvtárosa; *Hajnal Béla* kandidátus, a KSH Szabolcs-Szatmár-Bereg Megyei Igazgatóság igazgatója; *Marton Ádám* kandidátus, a KSH ny. osztályvezetője; *Szász Kálmán* kandidátus, a KSH Könyvtár és Dokumentációs Szolgálat tudományos kutatója; *Tűű Lászlóné*, a KSH ny. osztályvezetője; *Waffenschmidt Jánosné*, a KSH Budapesti és Pest Megyei Igazgatóság főigazgatója.

ISSN 0039 0690

Megjelenik havonta egyszer
Főszerkesztő: dr. Hunyadi László
Osztályvezető: Dobokayné Szabó Orsolya
Kiadja: a Központi Statisztikai Hivatal
A kiadásért felel: dr. Mellár Tamás
3249 – Akadémiai Nyomda
Martonvásár, 2001
Felelős vezető: Reisenleitner Lajos

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

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

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

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

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

TARTALOM

STATISZTIKAI ELEMZÉSEK

Gondolatok a felülvizsgálati politikáról. – <i>Hüttl Antónia – Pozsonyi Pál</i>	653
Az emberi élet értéke Magyarországon. – <i>Adorján Richárd</i>	669

MÓDSZERTANI TANULMÁNYOK

Az általánosított lineáris modell és biztosítási alkalmazásai. – <i>Roger Gray – Kovács Erzsébet</i>	689
A CHAID alapú döntési fák jellemzői. – <i>Hámori Gábor</i>	703

SZEMLE

A Kanadai Statisztikai Hivatal fejlesztései az ezredfordulón. – <i>Nádudvari Zoltán</i>	711
Magyar szakirodalom	
Hunyadi László: Statisztikai következtetésemélet közgazdászoknak. – <i>Oravecz Beatrix</i>	715

STATISZTIKAI HÍRADÓ

Személyi hírek	718
Szervezeti hírek – Közlemények	718

STATISZTIKAI IRODALMI FIGYELŐ

Külföldi statisztikai irodalom

Hand, D. J.: Az orvosi diagnosztikát szolgáló statisztikai módszerek. (<i>Szász Kálmán</i>)	720
Árindexkutatások a jövő évtizedekben. (<i>Marton Ádám</i>)	721
Új munkaerő-piaci trendek a közép-európai országokban. (<i>Hajnal Béla</i>)	722

Bruvoll, A. – Nyborg, K. – Halvorsen, B.: A háztartási hulladékok házilag szétválogatása. (Tűz Lászlóné)	725
Smirnov, S. N.: A regisztrált munkanélküliség prognosztizálásának módszere. (Bíró Júlia)	727
Schulz, R.: A világ népességének öregedése. (Waffenschmidt Jánosné)	728
Külföldi folyóiratszemle	730

*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!

GONDOLATOK A FELÜLVIZSGÁLATI POLITIKÁRÓL

HÜTTL ANTÓNIA – POZSONYI PÁL

A gazdaságstatisztika alapvető megújulásának világszerte végbemenő folyamata közepette talán túl korai a felülvizsgálati politika kérdéseivel foglalkozni. A felülvizsgálati politika a statisztikai késztermékek „marketing”-jéhez tartozik, és előbb kellene azzal foglalkozni, milyen statisztikai termékeket akarunk előállítani, majd ezt követően lehetne rendezni a köz-zététel módját és annak szabályozását. Napjainkban ugyan már valamennyire kialakult az az adathalmaz, amelyet a különböző mérvadó felhasználók a statisztikai szolgáltatótól igényelnek. Nagy vonásokban az adatok előállításának módszereit is ismertnek tételezhetjük fel. Sokkal kevésbé előrelátható azonban az adat-előállítási munkafolyamatban várható változások, az adatforrások, az elvárt megbízhatósági küszöbértékek, az adatgyűjtési technikák stb. változása. Ugyanakkor alapvetően ezek a tényezők befolyásolják az adatok elkészültének időrend-jét, és így azt is, hogy mikor keletkezik új információ. A statisztikai felülvizsgálat viszont nem más, mint az eddig ismert adatértékek módosítása az új információk felhasználásával. A kérdés tehát egyelőre még meglehetősen képlékeny, néhány alapelvben való hallgatóságos megállapodáson kívül nem beszélhetünk egységes nemzetközi gyakorlatról. Ebből következően az e tanulmányban tárgyaltak sem tekinthetők kiforrott álláspontnak. A problémák, dilemmák felvetésével elsősorban azt szeretnénk elérni, hogy felkeltsük a statisztikusok és a felhasználók érdeklődését a téma iránt.

TÁRGYSZÓ: Statisztikai felülvizsgálat. Nemzeti számlák.

A felhasználóknak gyakran okoz gondot, ha egyazon statisztikai fogalomhoz tartozóan több, többé-kevésbé különböző adattal találkozunk. Ilyenkor általában azt feltételezik, hogy pontatlanok a számítások. A statisztikai adatok változásának azonban ennél sokkal pozitívabb magyarázata van, mégpedig az, hogy a gazdaságról való ismeretünk az idő előrehaladtával egyre bővül és pontosabbá válik. Ezért a később közölt adatok minősége felülmúlja a korábban közöltet. Minden adatot úgy kell tekinteni, hogy az csupán az adatközlés időpontjáig feldolgozott információk felhasználásával készített legjobb becslés. Ha tudjuk, hogy az adatközlés után még újabb információk birtokába fogunk jutni, akkor először csak előzetesnek nevezzük az adatot, majd az később válik véglegessé. Helyesebben szólva, elméletileg sohasem beszélhetünk végleges statisztikai adatokról. A végleges jelző inkább csak arra utal, hogy ezt követően már nem érdemes újabb információkra várni, és ezért többé nem fog sor kerülni az adatok módosítására.

Statisztikai felülvizsgálatnak nevezzük a már nyilvánosságra hozott adatok visszamenőleges módosítását. A felülvizsgálati politika kifejezést abban az értelemben használjuk, hogy a hivatalos statisztikai szolgálat illetékes azt eldönteni, hogyan érvényesítse az adatmódosítás hatását a kapcsolódó statisztikákra, valamint hogyan tájékoztassuk minderről a felhasználókat.

Mind a felhasználók, mind a statisztikusok ódzkodnak az adatok felülvizsgálatától – a felhasználókat zavarja, hogy felül kell bírálniuk az eddig ismert adatokból levont következtetéseket, a statisztikusok számára pedig többletmunkát jelent az adatok ismételt átszámítása – tehát mindenki csupán szükséges rosszként hajlandó tudomást venni róla. Pontosabban fogalmazva: rossz, ha a felülvizsgálat hatására változik az adat, de jó, ha képesek vagyunk újabb információk, adatforrások birtokában többször átnézni és ellenőrizni az adatokat. Ideális esetben a felülvizsgált adat megegyezik az előzőleg közölttel. Ez teljes pontossággal sohasem következik be, viszont elvárható, hogy a statisztikai hibahatárokon belül legyen.

MILYEN ESETBEN KELL FELÜLVIZSGÁLNI AZ ADATOKAT?

A statisztikai felülvizsgálat négy különböző ok miatt nem kerülhető el:

- új információk beérkezése,
- a módszertan változása,
- adathibák javítása,
- adatértelmezési és egyéb technikai korrekciók esetén.

Újabb, pontosabb adatok beérkezése

A statisztikai felülvizsgálat elsődleges célja, hogy kezelje az adatközlés gyorsasága és megbízhatósága közötti ellentmondást. Csak az adatok rendszeres felülvizsgálata teszi lehetővé, hogy összeegyeztessük e két ellentétes szempontot. A gazdaságstatisztikák többségére jellemző, hogy a korai adatforrások pontatlanok és részlegesek. Később, az idő előrehaladtával egyre több, jobb és megbízhatóbb információ áll rendelkezésre. Az egymást követő időpontokban elérhető adatforrások között lehetnek explicit és implicit ellentmondások.

Vannak olyan esetek, amikor világosan, *explicit* módon kiderül, hogy a később rendelkezésre álló – jobbnak vélelmezett – adat eltér a korábban ismerttől. Például bizonyosnak vehető, hogy a vállalati számvitel mérlegzárás utáni adatai pontosabbak a vállalati mérleg összeállítás előtt szolgáltatott adatoknál. A korai időpontban kapott vállalati számviteli adatok tehát biztos, hogy csak előzeteseknek nevezhetők. Hasonlóképpen, a bővebb kört megfigyelő éves adatok jobbak, mint az évközi, kisebb mintával dolgozó felvételek. Az is előfordulhat, hogy az évközi adatnak is több, egyre pontosabb változata van.

Ide tartozik az is, hogy a censusokból (teljes körű összeírásokból) származó pontosabb, de csak több évvel a tárgyidőszak után elkészült adatok módosítják visszamenőleg az addig véglegesnek hitt ismereteinket a gazdaságról. Az évezredforduló környékén a teljes körű mezőgazdasági összeírás és a népszámlálás különösen időszerűvé teszi ezt a kérdést. Szinte bizonyosra vehető, hogy a mezőgazdasági összeírás eredményei kisebb-

nagyobb mértékben eltérnek azoktól az adatoktól, amelyeket a statisztika eddig a mezőgazdasági termelés szintjéről és összetételéről közölt. A népszámlálás eredményei áttételesebben befolyásolják a gazdasági adatokat. Két kapcsolódási területet azonban már előre lehet látni. A népszámlálásnak a foglalkoztatás szintjére és összetételére vonatkozó adataiból következtetni lehet a gazdasági tevékenységek és a keletkezett jövedelmek azon részére, amelyet a szürke, informális, rejtett jelzőkkel szoktunk specifikálni. Bármi legyen is az eredmény, a statisztika azt nem hagyhatja figyelmen kívül. Mindenképpen állást kell foglalnia abban a kérdésben, hogy az eddig különböző részleges forrásokból becsült adatokat korrigálja-e a népszámlálás adatainak ismeretében. Hasonlóképpen többletinformációkhoz juthatunk a népszámlálásból a lakásállomány összetételéről, és ebből számítva változhat a lakásszolgáltatás értékbecslése. Fontos hangsúlyozni, hogy az eltérés egyik esetben sem a statisztikai munka gyengeségét jelzi. Ha nem számítanánk arra, hogy a censzusok jobb, pontosabb képet adnak az eddigiéknél, akkor feleslegesek lennének. Ezért arra is fel kell készülni, hogyan érdemes a bizonyosra vehető korrekciókat több évre visszamenően átvezetni a már közzétett statisztikákon.

Akkor beszélhetünk *implicit ellentmondásról*, ha két statisztikai adat nem teljesen ugyanarra a jelenségre vonatkozik, de ez a körülmény önmagában nem magyarázhatja az adatok közötti különbség mértékét. Példa erre a háztartások fogyasztásának becslése: a kiskereskedelmi forgalom nagy részét a háztartások vásárlása adja, de a forgalom nem ismert részét a közösségi vásárlások vagy a kisvállalkozások beszerzései teszik ki. Külön probléma, hogy nem lehet elkülöníteni a hazai (rezidens) háztartások vásárlásait a külföldi turistákétól. Ez az egyik adatforrás. Emellett létezik felvétel a magyar háztartások kiadásainak megfigyelésére. Valamely termékcsoport kiskereskedelmi forgalomban kimutatott keresletének nem kell feltétlenül megegyeznie a háztartás-statisztika kiadási adatával, de az eltéréseknek bizonyos elfogadható korlátokon belül kell maradniuk. Termék- és szolgáltatáscsoportoktól függően más-más mértékű eltérést tekinthetünk még elfogadhatónak. Nem tartósított élelmiszereket a kiskereskedelmi forgalomban lényegében csak a hazai háztartások vásárolnak, ez esetben a két adatforrásból származó adatnak egyeznie kellene. A vendéglátóhelyek forgalmának jelentős részét viszont a turisták fogyasztása teszi ki. Így ennél jóval alacsonyabb adatot várunk a háztartási kiadási felvételtől. (Igaz ugyan, hogy a magyarok is járnak külföldre, és az ottani kiadásait a háztartás-statisztika is elszámolja, de azt nem a vendéglátó-ipari, hanem az idegenforgalmi kiadások között.) A személygépkocsik kiskereskedelmi forgalmazásában igen magasnak valószínűsíthetjük a vállalkozások beszerzéseinek arányát (más kérdés, hogy a gépkocsit a vállalkozás tulajdonosa saját személyes céljaira is használja), így a háztartás-statisztika itt is alacsonyabb értéket mutat. Minden egyes termékcsoporthoz tehát külön mérlegelni kell, a kiskereskedelmi forgalom és a háztartás-statisztikai felvétel közül melyiket vagy a kettőből hogyan átlagolva származtathatjuk a nemzeti számlában a háztartások fogyasztásának termékszerkezetét. Ráadásul mindkét statisztikának van egyszerűsített évközi és részletes éves változata, így a háztartások fogyasztási szerkezte többször is felülvizsgálatra szorul.

Tulajdonképpen a felülvizsgálatnak ehhez a típusához tartozik az az eset is, amikor felfrissítik az ár- vagy volumenindex-számításhoz használt súlyokat. Minél időszerűbbek a súlyok, annál jobban tükrözik a gazdaságban a megfigyelés idején érvényes érték- vagy szerkezeti arányokat. Bár ebből az is következik, hogy minél ritkábban változtatják a sú

lyokat, annál elavultabb a használt struktúra, a túl gyakori váltásnak is vannak hátrányai, ahogy azt az indexszámításból ismerjük.¹

A statisztikai módszertan változása

Ideális esetben a hivatalos statisztika módszertana örök és állandó. Ez biztosítaná az adatok tökéletes összehasonlíthatóságát térben és időben. A módszertan változatlanságának azonban az adekvátság elvesztése lehet az ára, azaz, hogy a statisztikában nem követjük a megfigyelendő gazdasági jelenségek természetének változását. Képzelnék csak el, mit mutatna a mai gazdaságban a kilencvenes évek előtt alkalmazott ágazati osztályozás, amelyben a közigazgatás és a pénzügyi szolgáltatások ugyanazon nemzetgazdasági ágba tartoztak. A kilencvenes években végrehajtott módszertani reform talán kiugró példa, de illúzió lenne azt remélni, hogy a jövőben nem számíthatunk további jelentős változásokra. Nemcsak az EU már ismert követelményeinek az átvétele miatt, hanem azért is, mert a követelmények is változnak, követve részben a világgazdaság, részben az európai integráció erősödését. Ma még szinte beláthatatlan, milyen hatásai lesznek a többnyire közhelyként emlegetett globalizációnak és információs forradalomnak a gazdaságstatisztikára.

Széles értelemben módszertani változásnak tekinthetjük a fogalmak, osztályozások, a megfigyelt sokaság és a mintavételi technikák stb. változásának, az új adatgyűjtések bevezetésének az adatokra gyakorolt hatását. Azt is a módszertani változás részének tekinthetjük, amikor módosítjuk azokat a heurisztikus vagy algoritmizált összefüggéseket, amelyek révén az évközi konjunktúramutatóktól eljutunk a negyedéves nemzeti számlákban megjelenő adatokig. A nemzeti számlák annál jobb minőségűek, minél több adatforrást tudnak egyetlen konzisztens keretbe illeszteni. Így attól függően, hogy milyen új évközi alapstatisztika vagy más, például államigazgatási nyilvántartás válik hozzáférhetővé, más és más becslési eljárásokat célszerű alkalmazni. Mivel azonban a negyedéves nemzeti számlák egyelőre nagyrészt heurisztikus becslési módszerekkel készülnek, és ez így is lesz mindaddig, amíg nincs kellően hosszú idősor a paraméterek statisztikai úton történő becsléséhez, a módszertani változtatás ténye sok esetben még a becslést végzők számára sem válik nyilvánvalóvá. Ez az egyik olyan terület, ahol – ahogy a bevezetőben hangsúlyoztuk – még nem érdemes hosszabb távra előre rögzíteni a gazdaságstatisztika felülvizsgálati politikáját. (A kérdésre a későbbiekben, a rendszerváltás statisztikájával kapcsolatban még visszatérünk.)

Az eddigiekből következően azt is mondhatjuk, hogy a statisztikai módszertan folyamatosan változik. Folyamatosan azonban képtelenség a módszertani változásokat a statisztika egészén, konzisztens módon átvezetni, beleértve az idősorok visszamenőleges korrekcióját. Ebben a tekintetben a felülvizsgálati politika művészete, hogy hogyan tudja megkülönböztetni a lényegtelen kis és a lényeges nagy módszertani változtatásokat. Az előbbiek csendben elrendezhetők, legfeljebb lábjegyzetben megemlítve, hogy változás következett be, és a változást nem érdemes visszamenőleg végigvezetni sem az idősoron, sem az adott statisztikából származtatott további adatokon. A lényeges, nagy módszertani változásokat viszont célszerű bizonyos előre kiválasztott évre koncentrálva, „csomagban”

¹ A változó súlyú láncindexek egyik legnagyobb hiányossága a gazdaságstatisztikai alkalmazás esetén, hogy nem biztosítják az additivitást, vagyis azt, hogy az idősorban a részindexek súlyozott átlaga kiadja a főindexet.

bevezetni. Az is előnyös, ha a kiválasztott év egyben az árbázisváltás éve, az az év, amikor változnak a volumenindexek számításához használt súlyok. A változtatás ténye mindenképpen szembevetendő, hiszen erről az évről két adatsor készül, a régi és az új áron. Nincs is szükség külön felhívni a felhasználók figyelmét a változásra. A módszertani változtatás és a súlyozás évének váltása összekapcsolható ugyan, de nem ugyanúgy kezelendő. A súlyok változtatása esetén elegendő egyetlen évre két adatsort készíteni, értelmetlen viszont az idősort visszavezetni. (Ebben az esetben ugyanis nem tennénk mást, mint valamilyen jövőbeli értékarányokkal súlyoznánk a volumenek változását. Ez a megközelítés ellentmond a tényekkel dolgozó statisztika alapelveinek.) Ezzel szemben a módszertani változásokat ajánlatos minél hosszabb időszakra visszavezetni. A felhasználók homogén idősorokat igényelnek. Az elmondottakból az is következik, hogy bár célszerű összekapcsolni a módszertani váltást és a volumenindexek súlyozásának változtatását, de nem szabad a két korrekciót egyetlen adatsorba összemosni. A módszertani változás hatását elkülönítve is ki kell mutatni, sőt, lehetőleg több évre visszavezetve.²

Adathiba miatti utólagos javítás

Jól működő statisztikai rendszerben – az adatközlést követően észrevett – adathibának nem szabadna előfordulnia. A világszerte most kiépülő statisztikai minőségbiztosítási rendszerek várhatóan jelentősen csökkentik az adathibák utólagos felfedezésének a veszélyét. Amennyiben azonban adatközlés után mégis kiderül, hogy az adat hibás, a megoldást egyedi módon kell megkeresni. A jelentős hibák eltitkolása jobban rontja a statisztikai szolgálat tekintélyét, mint a hiba jól kommunikált beismerése. Azt is mérlegelni kell, mikor van szükség a hibát soron kívül, a következő tájékoztatási időpontot be nem várva korrigálni, és mikor megengedett elhalasztani a közzétételt. Minden esetben a felhasználók szempontjai az elsődlegesek.

Adatértelmezési és egyéb technikai korrekciók

Hagyományos megközelítésben a hivatalos statisztikai szolgálat arra szorítkozik, hogy a tényadatokat eredeti formájában tegye közzé a felhasználókra bízva azok statisztikai eszközökkel való további elemzését. Az utóbbi években világszerte megfigyelhető, hogy a statisztikai szolgálat is alkalmaz olyan technikákat, és közöl olyan hivatalos adatokat, amelyek a megfigyelt adatok statisztikai elemzési módszerekkel való becsülésével, modellezéséből származnak. A statisztikusok sok esetben felkészültebbek e technikák alkalmazásában, mint az átlagos felhasználó, akit csak az érdekel, mi következik az adatokból a jövőre nézve, és kevéssé az, hogyan készül maga az adat. A többlétszolgáltatásokkal a statisztikai szolgálat az egyre bővülő, de a statisztikai módszerek alkalmazásában nem jártas felhasználói kör igényeit akarja kielégíteni. Az információs technológia elterjedése is közrejátszik abban, hogy elmosódtak a határok a valóság megfigyelése és a megfigyelés eredményének további kezelése között. Tipikusan ilyenek a szezonálisan

² A közelmúltban ezt a feladatot kellett megoldani az EU-tagországoknak: 1999-ben az 1995. évtől kezdve kellett átszámolni a nemzeti számlákat az ESA95 módszertana szerint. A minimumkövetelmény négyéves idősor volt, de – ha nem is a nemzeti számlák teljes részletezettségében, hanem összevontabb szinten – a jó statisztikai gyakorlatot folytató országok 20-40 éves vagy akár még hosszabb idősort is homogenizáltak.

kijáratott évközi adatok, trendek. Ide tartozik például a modellel számított maginfláció is. Vannak országok, ahol a negyedéves GDP is ökonometriai modellel készül.

Az idősor-modellezésben az idősor megfigyelt értékét egyelemű mintának tekintjük, és ebből következtetünk az elméleti függvényre. Ha a következő hónap, negyedév, év adatainak beérkezésekor az idősor meghosszabbodik, ezt úgy vesszük, mintha új mintához jutnánk. Az új mintából más becslést kaphatunk az egész modellezett időszakra. Ebből következően a múlt adatai is – minden új megfigyelési időpontban – változhatnak. A felülvizsgálati politika során arról kell döntenünk, milyen gyakran célszerű újrabecslülni az elméleti függvényt.

A technikai korrekciók közé soroljuk a negyedéves és éves nemzeti számlák egymásra épülése miatt előforduló azon esetet, amikor a bázisidőszak felülvizsgálata miatt változnak az azt követő negyedéves időszakok adatai. Ugyanis ebben az esetben nem a megfigyelt időszakra vonatkozó adatforrások módosulnak, pontosodnak, csupán arról van szó, hogy a viszonyítási alap (a volumenindexek nevezője) változik.

MILYEN ADATKÖRÖKRE TERJEDJEN KI A FELÜLVIZSGÁLAT?

Első megközelítésben a válasz kézenfekvőnek látszik. Minél több adatkört tud aktualizálni, annál jobb és egységesebb a statisztikai rendszer. Legfeljebb a munkaerőforrás-igényessége szabhat korlátot. Jobban megfontolva azonban differenciáltabb választ kell adnunk. Felülvizsgálatra gyakran csak akkor kerül sor, amikor az adat információtartalma már elévül, mindenkit a következő időszakra vonatkozó adat érdekel. Ez bizonyos mértékben így van az éves nemzeti számláknál is, a rövid távú konjunktúramutatók esetén magától értetődő az érdeklődés gyors lanygulása. Ilyen körülmények között a felülvizsgált adat csak akkor fontos, ha belőle következtetni lehet a jövőre, arra, hogy a következő időszakban a felülvizsgált adat hogyan fog eltérni az előzetestől, és ezzel – gondolatban – már előre pontosítani lehet az előzetes adatot. Ha a felülvizsgálat erre – a jelenség erősen ingadozó természete miatt – nem alkalmas, akkor megfontolandó, érdemes-e az évközi konjunktúramutatókat felülvizsgálni. Mivel év közben minden jelenségre rendszerint csak egy megfigyelést végzünk, felülvizsgálatra legkorábban az éves adatok feldolgozását követően kerülhet sor, több mint egy évet kellene várni a pontosabb adatra. Ugyanakkor a priori feltételezhetjük, hogy az éves adat jobb, mint az évközi. Az éves adatgyűjtés alapja ugyanis nagyobb minta, több idő áll rendelkezésre az adatok begyűjtésére és ellenőrzésére. Az évközi adatok esetén viszont a gyorsaság fontosabb, mint a pontosság.

Az évközi adatok felülvizsgálata

Az eddigiek alapján az évközi és az éves alapstatisztikák között a következő logikai kapcsolat képzelhető el.

Évközi és éves indexek. Évközben fontosabb a változások nyomon követése, mint a szintek ismerete. A megfigyelésből kimaradt körre – egyéb ismeret híján – szükségképpen azt tételezzük fel, hogy az ugyanúgy változik, mint a megfigyelt. Ezt a feltevést ellenőrzi az éves felvétel. Ezért amennyiben lényegesen eltérnek az éves és évközi változási ütemek, akkor az évközi adatgyűjtésen, például a minta nagyságán kell változtatni.

Amennyiben az eltérés nem jelentős, akkor viszont nincs sok értelme a visszamenőleges korrekciónak.

A szintekre vonatkozó évközi és éves adatok. Kifejezetten megtévesztő lehet gazdasági jelenségek szintjéről részleges adatokat közölni. Ezért, ha a statisztikai szolgálat mégis úgy dönt, hogy évközben a szintek alakulásáról abszolút számokban tájékoztat, akkor három megoldás közül választhat.

– Elsősorban arra kell törekednie, hogy a közölt adatba eleve becsülje a nem megfigyelt rész értékét, felhasználva például azt az ismeretet, mekkora volt a korábbi években az éves és az évközi adatok különbsége. Ezzel a módszerrel például kellő megbízhatósággal meg lehetne állapítani a megfigyelésből kimaradó kis gazdasági szervezetek dolgozói létszámát, és évközben is közölhető lenne a teljes foglalkoztatotti létszám.

– Ha az előbbi módszer alkalmazására nincs mód a hiányzó rész előre kiszámíthatatlanul változó természete miatt, akkor a fogalom elnevezésével jelölni kell, hogy a jelenség milyen részére vonatkozik az érték. Például évközben csak a rendszeres pénzbeni bérek és keresetek összegét célszerű megadni, ez a mintából, vélhetően, jól teljeskörűsíthető. Félő, hogy sokkal jobban torzítana az a mutató, amelybe „szakértői becsléssel” a természetbeni, valamint a nem rendszeres pénzbeni keresetelemeket is beleszámítanák.

– Tudomásul kell venni, hogy vannak olyan jelenségek, amelyekre évközben nem lehet kellően megbízható értékadatokat közölni. Ilyen például a vállalati nyereség, amelyet évközben sem megfigyelni, sem a múlt alapján megközelíteni nem lehet. A vállalati nyereség maradék, amelynek összege csak a gazdálkodás egészének ismeretében számítható ki. Hasonlóképpen bizonytalan a háztartási jövedelmeknek a háztartási megfigyelésekre építő évközi megfigyelése, különösen azért, mert a statisztikának nem a ténylegesen befolyt, hanem eredmény szemléletben a háztartásnak „járó” jövedelmet kellene számítani. Tulajdonképpen a mezőgazdaságban megtermelt hozzáadott értéket sem lehet évközben, a termelési ciklus lezárása előtt meghatározni.

Összefoglalva tehát több érv szól amellett, hogy az évközi alapstatisztikák esetén felesleges a tárgyidőszakban erőfeszítéseket tenni a felülvizsgálatra. Az évközi gyors tájékozódáshoz előnyben kell részesíteni a változást mutató indexeket, és gondosan mérlegendő, milyen gazdasági jelenségek szintjét lehet megbecsülni kellő megbízhatósággal.

Más a helyzet a negyedéves és az éves nemzeti számlák viszonyában. A nemzeti számlák elsősorban a jelenségek konzisztens, átfogó leírására törekuszenek. Ebből következően év közben is vállalkozniuk kell arra, hogy valamilyen módon kezeljék még a rosszul becsülhető folyamatokat is. Nem mellőzhetők a szintekre vonatkozó számítások sem, sőt becslést kell adni ezek árváltozására is. Így kizárt, hogy a negyedéves nemzeti számlák ne szorulnának akár többször is jelentős felülvizsgálatra. Éves szinten több és jobb minőségű statisztika áll rendelkezésre, mint év közben. Ezért az éves nemzeti számlák felülírják a negyedéveseket. Sőt a tökéletes összhang biztosítása érdekében ahány változatban az éves számlák készülnek, annyiszor pontosítják ugyanazon év negyedéves adatait.

Az éves adatok felülvizsgálata

Az éves adatok esetén két kérdés vár tisztázásra. Az első, hogy mi a teendő az éves alapstatisztikákkal az éves nemzeti számlák elkészülte után. Érdemes-e arra törekedni, hogy a „történelmi idősorban” végül is csak egyetlen adat létezzen például az ipari termelés értékére. A felhasználói igények sokrétűségét szem előtt tartva a válasz nem egyértelmű. A éves szakstatisztikáknak az az elsődleges célja, hogy bemutassák a részletes strukturális összefüggéseket: terméksoros adatokat a termelés, az értékesítés, a külkeres

kedelem alakulásáról, részletes ágazati bontásban a termelési szerkezetet, bizonyos speciális jelenségek kiemelt megfigyelését. Ilyen, a szokásos ágazati szerkezetből kilógó jelenség többek között a kutatás–fejlesztés, a háztartások jövedelmének társadalmi rétegek szerinti vizsgálata. A nemzeti számlák feladata, hogy összefüggően ábrázolják az újratermelési körforgás egészét, bemutassák az állományok és a folyamatok közötti kapcsolatot a nemzetgazdaság egészében. A nemzeti számlákban nem a részletek önmagukban, hanem a részek és az egész összefüggése jelenti a fontos információt.

Ebből következően az éves szakstatisztika és a nemzeti számlák megfelelő adatának nem kell egy az egyben megfelelnie egymásnak. Sőt kifejezetten tisztítja a képet, ha például a szakstatisztika termelési érték (kibocsátás) adatában nem jelenik meg a kiegészítő tevékenységként végzett saját számlás beruházás vagy a saját célra történő K+F értéke. A nemzeti számlákban a teljes körű elszámoláshoz viszont számba kell venni ezen tevékenységeket is. A munkaügyi statisztikában is előfordulhat, hogy a természetbeni munkajövedelmek bizonyos tételeit nem tudjuk felbontani a különböző foglalkoztatotti rétegek (nők–férfiak, szakmák) között. Ezért elfogadható az a megoldás, ha a természetbeni jövedelmek értéke kimarad a részletes kereseti adatokból. A nemzeti számlákban ugyanakkor el kell számolni a munkavégzésből származó összes – akár pénzbani, akár természetbeni – jövedelemmel.

Bizonyos eltérések tehát megengedettek, gondoskodni kell azonban a felhasználók megfelelő tájékoztatásáról. Ne csak az eltérés tényével szembesüljenek, hanem azt is megismerhessék, hogyan vezethető le a szakstatisztikából a nemzeti számlában megjelenő adat.³ A nemzeti számlák összeállítása során azonban kiderülhetnek olyan adathibák is, amelyek korrekcióját a szakstatisztikákon is át kell vezetni. Ebből következően az éves szakstatisztikák véglegesítésére is csak a nemzeti számlák véglegesítésével együtt kerülhetne sor. Nem számszaki egyezőséget kell tehát elérni, hanem az eltérések között a „híd”-tételeket kell kimutatni.

Az éves nemzeti számlák esetén a felülvizsgálat kötelező mindaddig, amíg vélelmezhető, hogy a később még befutó információk már érdemben nem javítják az adatokat. Ennek mérlegelése elsősorban úgy merül fel, hogy érdemes-e bevárni a censzusok feldolgozását vagy az adóbevallások akár több évet is igénybe vevő ellenőrzési folyamatának lezárását.

Egyéb okokra visszavezethető felülvizsgálati politika

Mint már szó volt róla, a jelentősebb módszertani változásokat célszerű egy előre kijelölt fordulóponthoz kötve „csomag”-ban bevezetni. A nemzetközi gyakorlatban az 5-tel és a 0-val végződő évekhez szokták időzíteni a jelentősebb módszertani változásokat, együtt azzal, hogy ezeket az éveket teszik a változatlan áras számítások árbázisává. A partneri viszony ápolása, a statisztika átláthatósága érdekében jó, ha a felhasználók nem csak utólag, az adatok megjelenésekor szereznek tudomást a módszertan módosításáról, hanem erre már jó előre felkészülhetnek, sőt lehetőség szerint belevonják őket is a szakmai előkészítésbe.

³ Szembetűnően felhívja a figyelmet az eltérésekre, ha a nem egybevágó fogalmakhoz eltérő elnevezések társulnak. Ennek érdekében a beruházási statisztikában a szakstatisztika a beruházás, a nemzeti számlákban az állóeszköz-felhalmozás kifejezést használja.

Adathiba észlelése esetén alapelveként kimondhatjuk, hogy a hiba mértékétől függően vagy azonnal javítani kell, vagy ezzel megvárható a következő rendszeres felülvizsgálat időpontja. Minthogy kivételes helyzetről van szó, egyedi mérlegelés alapján kell gondoskodni arról, hogyan történjék a felhasználók tájékoztatása. Az adatértelmezési és egyéb technikai korrekciókról az átlag-felhasználónak viszont elég annyit tudnia, hogy előfordulhat, de nem kell minden egyes alkalommal külön foglalkozni vele. Természetesen kivéve azt az esetet, amikor kirívóan nagy az eltérés mértéke.

A MAGYAR STATISZTIKAI GYAKORLAT

A magyar statisztikai szolgálatban az újabb adatok beérkezését, az adatok javulását követő rendszeres és tervezhető felülvizsgálat eddig lényegében a nemzeti számlákra korlátozódott. Bár olyan statisztikák is léteznek, amelyek több, előzetes és végleges változatban is ismertek, a felülvizsgálat csupán azt jelenti, hogy az első adatközlés nem várja meg az adatellenőrzés befejeztét, hanem néhány összevont jelzőszám félig kész állapotban is nyilvánosságra kerül. Ez a tájékoztatási rendszer például a nemzetközi fizetési mérleg vagy a havi ipari termelés esetén. Az alapstatisztikák közül eddig az egyetlen kivétel a kiskereskedelmi forgalom mutatója. Ennél a kiskereskedelmi regiszter nyilvántartásának késedelme és a bolthálózat változékonysága miatt nem kerülhet meg a regiszter aktualizálása után a minta újbóli teljeskörűsítése.

Egyéb statisztikák esetén az éves és évközi szakstatisztikák is önálló életet élnek, vagy a 12 hónap összesen adata különbözik az éves adattól, vagy a havi adatok összesen jele az éves értéket. Tudatos felülvizsgálati politikáról egyik esetben sem beszélhetünk.

Az éves nemzeti számlák esetén elkerülhetetlen a rendszeres felülvizsgálat. E számlák lényegében a makrogazdaság helyzetével kapcsolatba hozható valamennyi statisztikát összefoglalják. A legkorábbi adat a konjunktúramutatók tizenkét hónapra összesített értéke, míg a leginkább időigényes a forrásfelhasználási táblák összeállítása. Eközben több mint egy, esetleg két év telik el. Ez idő alatt a nemzeti számlák több változata kerül közzétételre.

A múltban (2000-ig) az éves nemzeti számlák három változatban készültek:

– a GDP fő összetevőire 3-4 hónappal a tárgyév lezárása után „szakértői becslés” készült az addig rendelkezésre álló évközi statisztikák összesítésével; az adatközlésre nem előre megadott időpontban került sor, hanem az megvárta a megelőző évi GDP-adat véglegesítését, így egyben a bázisévi változás hatását, mint technikai korrekciót is figyelembe lehetett venni;

– az éves előzetes nemzeti számlák körülbelül 10 hónappal a tárgyévét követően, október hónap folyamán jelentek meg, lényegében az adóbevallások gyors feldolgozását követően; ezek az előzetes adatok részlegesen is voltak, nem vizsgálták a teljes kört, hanem a nemzetgazdasági ágak és szektorok szerinti hozzáadott érték alakulására (folyó és változatlan áron), a GDP felhasználásának részleteire, valamint a háztartások jövedelmének és fogyasztásának főbb tételeire korlátozódtak;

– a végleges adatok a tárgyévét követő körülbelül 16 hónap után készültek el; az októberi előzetes és a végleges adatok lényegében ugyanarra az adatforrásra támaszkodtak, az adatok véglegesítése során nyílt lehetőség az adatforrásokban lévő hibák kijavítására, bizonyos részletek kidolgozására és az összefüggésvizsgálatokra (a végleges adatok megelőzték a termelés részletes forrásfelhasználási összefüggéseinek vizsgálatát, az ágazati kapcsolatok struktúrája utólag került az előre adott keretbe).

A negyedéves GDP-becslések bevezetése az éves adatok közzéadási rendjét mindaddig nem befolyásolta, amíg negyedéves szinten csupán volumenindexek készültek, és azok nem lettek összeillesztve az éves adatokkal.

Az éves termelési és munkaügyi szakstatisztikák nem épültek be a nemzeti számlákba. A nemzeti számlákban a bruttó kibocsátás és a hozzáadott érték adatai alapvetően nem statisztikai adatforrásokból származtak, hanem a társasági adóbevallásokból és a költségvetési intézmények beszámolóiból. Készültek ad hoc összehasonlítások az éves ipari termelés kétféle volumenindexe között kimutatott – esetenként jelentős – eltérés feltárására, ebből további módszertani tanulságok származtak, korrekcióra azonban emiatt nem került sor.

Ez a helyzet 2000 áprilisában a negyedéves nemzeti számlák esetében megváltozott. A Központi Statisztikai Hivatal attól kezdve már nemcsak a volumenindexeket közli negyedévente, hanem a GDP felhasználási összetevőinek értékadatait is. Azóta a GDP negyedéves folyó és változatlan áras felülvizsgált adatai az éves adatokkal harmonizáltan jelennek meg.

A felhasználó, ha tudomásul is veszi a felülvizsgálatok tényét és azt, hogy ez is az ő érdekében történik, nagyon nehezen tud eligazodni az eltérő adatok között. Ahány kiadványt elővesz, annyi adattal találkozhat ugyanarra a jelenségre vonatkozóan. A jövőben, az internetes tájékoztatás bővülésével, még áttekinthetlenebbé válik a helyzet. Használati utasításként az adható, hogy mindig a legfrissebb adat a legjobb. Ehhez még azt kell hozzátenni, hogy mikor várható az új adatváltozat.

A rendszeres és tervezett felülvizsgálatok tájékoztatási határideje egy-két évre előre mindig megadható. Változást ugyanis új adatforrások belépése idézhet elő, ennek átfutási ideje pedig több évet vesz igénybe. Ennél hosszabb távra azonban nehéz előre látni, mert mint a bevezetőben említettük, a statisztikák előállítási folyamata világszerte változóban van. További támpontot adnak a tájékoztatási időpontok kitűzéséhez az EUROSTAT által megszabott adatátadási határidők. A nemzeti számlák esetén például a tárgyévet követő 3., 9., 12. és 36. hónapban kell új, egyre részletesebb adatokat átadni. A harmadik év végére kell elkészíteni a forrás-felhasználási táblákkal alátámasztott és a számlák teljes sorozatát tartalmazó végleges dokumentumot.

Az uniós követelményekhez két szempontból is célszerű alkalmazkodni. A nemzeti hivatalok a határidőt megelőzően nyilvánosságra hozhatják az adatokat, de a határidő elég feszes, főleg az éves számlákra vonatkozó 9. és 12. hónap. Így nem valószínű, hogy sokkal korábban el lehet készülni. A végleges éves adatok átadására rendelkezésre álló három év hosszúnak tűnik, de ne felejtsük el, a jó nemzeti számlák jellemzőjének is nevezhetjük, hogy még évekre visszanyúlva is tudnak új információk alapján újat mondani. Nem célszerű tehát túl korán lezárni a felülvizsgálat folyamatát. Hasonló a helyzet a részletes éves alapstatisztikák esetén, itt is az javítaná a statisztika minőségét, ha érdemi munkával lehetne kitölteni a rendelkezésre álló 18 hónapot.

A negyedéves számlák gyakori felülvizsgálatát az új információk két oldalról is kényszerítik. Egyfelől a rövid távú évközi statisztikák változása mindig befolyásolja az arra épülő negyedéves nemzetiszámla-adatokat. Másfelől az éves számlák új változata visszamenőleg módosítja az addigi puhább adatforrásokra alapozott negyedéves adatokat. Technikai korrekciót jelent, hogy az éves számlák változása, mint a bázis változása, módosíthatja az azt követő összes negyedéves számla volumenmutatóját.

A felülvizsgálati politikát tovább bonyolítja az a körülmény, hogy megkezdődött a negyedéves ún. „flash estimate” közreadása. Ez azt jelenti, hogy a negyedéves szerkezeti statisztikák alapján körülbelül 60 nap alatt összeállításra kerül a hozzáadott érték ágazati volumenindexe. Ebben az esetben viszont a felhasználóknak meg kell barátkozniuk a negyedéves GDP még egy további, érdemi változatával.

MÓDSZERTANI FELÜLVIZSGÁLAT AZ EURÓPAI UNIÓBAN

Hosszú évekig tartott, amíg átdolgozták az Európai Számlák Rendszerét.⁴ Az új változat, az ESA95 az Európai Unió tagállamai számára kötelező jelleggel írja elő a nemzeti számlák készítésének módszereit és az adatközlési határidőket. A tagállamokban – hat évvel az új metodika elfogadása után – még jelenleg is tart a nemzeti számláknak az új előírásokhoz igazítása. Az ESA95 alkalmazása egyben lehetőséget adott arra, hogy a tagállamok rendszerszerűen újragondolják és korszerűsítsék a nemzeti számlák adatforrásait és módszereit. A hozzáigazítás arra is alkalmat kínált, hogy a nemzeti számlák statisztikájában új minőséget hozzanak létre. Az ESA95 ugyanis nem egyszerűen a korábbi nemzeti számla módszertan kisebb korrigálását jelenti, hanem átfogó koncepcionális változásokat vezetett be, válaszolt azokra az új gazdasági jelenségekre, amelyek az 1980–1990-es években váltak meghatározóvá. Az egyes tagországok felülvizsgálati programjai időben eltérnek, de összekapcsolja őket két kötelező határidő:⁵ az egyik az 1999. év a kiemelt mutatók, a másik a 2002. év a számlák teljes összeállítására. Az is igaz ugyan, hogy az ESA95-jogszabály⁶ alapján készült adatközlési program alól néhány tagország felmentést kért, mind az első adatközlésre, mind a végső határidőre vonatkozóan. A többéves felkészülési, átállási időszak és a derogációk léte is mutatja, mekkora munkáról van szó.

Az EUROSTAT 1999 júliusában nyilvánosságra hozta a felülvizsgálat első szakaszának eredményeit.⁷ A 15 tagország – Ausztria, Görögország és Luxemburg kivételével –, az 1995. bázis évre vonatkozóan készítette el az új számításokat. A közölt adatok azt mutatják, hogy az ESA95-re való áttérés összességében mérsékelt hatást gyakorolt a GDP és a GNP (GNI) nagyságára, ha azokat a korábbi, ESA79 szerint készült és korábban már közölt adatokhoz viszonyítjuk. E megállapítás igaz mind az értéki (színvonal-), mind a volumenindex-mutatókra. A változtatás hatása az egyes GDP-elemeket illetően már erőteljesebb, a hatások az összevont GDP szintjén egyenlítik ki egymást. Általánosságban megállapítható, hogy a tagállamok többségében az új módszertan szerint becslült GDP nagyobb, mint a korábbi. A változás a tagállamok összességének (EU15) szintjén mintegy +2,0 százalékot tesz ki.

Az 1. tábla adatai szerint, az egyes tagországok szintjén a változások jelentősen eltérnek: míg a módszertani váltás hatására a GDP szintje Írországból mindössze +0,2 százalékkal emelkedett, Dániában viszont +6,4 százalékkal. A változásban a szigorúan vett módszertani okok mellett közrejátszott:

⁴ European System of Accounts (ESA95).

⁵ A megállapodás az Európai Gazdaság Tanácsa tagjaira (EU-tagállamok és Norvégia, Svájc, Izland) vonatkozik.

⁶ Council Regulation 2223/96 (25 June 1996).

⁷ Implementation of ESA95 Major progress by statisticians to improve the National Accounts (EUROSTAT New Release, No 69/99, 16 July 1999). További információk az EUROSTAT 1999. júniusi „Meeting of the Working Party on National Accounts” ülésére benyújtott „Evaluation of the first transmission of ESA95 data”, illetve az „Impact of the transition from ESA79 to ESA95” jelentéseiben állnak rendelkezésre.

- az új, jobb statisztikai források felhasználása,
- a szokásos, rendszeres felülvizsgálatok hatása,
- a makrogazdasági összefüggések szélesebb körű és megbízhatóbb megfigyelése.

1. tábla

Az ESA95 és az ESA79 adatai közötti különbség százalékban, 1995

Terület	Bruttó hazai termék (GDP)			Végső fogyasztási kiadás	Bruttó felhalmozás	Export	Import
	összesen	koncepcionális	statisztikai forrás és egyéb				
Eurózána	+ 1,9	:	:	+ 1,1	+ 7,2	+ 2,5	+ 4,1
EU15	+ 2,0	:	:	+ 1,1	+ 7,2	+ 2,2	+ 3,4
Belgium	+ 0,8	+ 1,6	- 0,8	- 2,0	+ 14,3	+ 5,7	+ 6,6
Dánia	+ 6,4	+ 4,1	+ 2,3	+ 4,3	+ 17,1	+ 5,3	+ 5,8
Németország	+ 2,3	+ 1,1	+ 1,2	+ 1,3	+ 6,4	+ 5,1	+ 5,5
Görögország	:	:	:	:	:	:	:
Spanyolország	+ 4,4	+ 1,5	+ 2,9	+ 3,2	+ 10,3	- 0,6	+ 0,6
Franciaország	+ 1,2	+ 0,2	+ 1,0	+ 1,2	+ 6,7	- 3,2	+ 1,0
Írország	+ 0,2	- 3,0	+ 3,2	+ 5,1	+ 5,8	+ 0,3	+ 6,9
Olaszország	+ 0,9	+ 1,7	- 0,8	- 0,4	+ 7,0	- 1,7	- 1,5
Luxembourg	:	:	:	:	:	:	:
Hollandia	+ 4,1	+ 3,3	+ 0,8	+ 2,7	+ 13,7	+ 12,8	+ 15,8
Ausztria	:	:	:	:	:	:	:
Portugália*	+ 1,9	:	:	+ 0,3	+ 3,8	+ 0,9	- 1,2
Finnország*	+ 2,1	:	:	+ 0,5	+ 12,9	+ 1,0	+ 2,2
Svédország	+ 3,4	+ 2 – +2,5	+ 1 – +1,5	+ 1,8	+ 11,1	+ 1,6	+ 1,1
Egyesült Királyság	+ 1,6	+ 0,8	+ 0,8	+ 1,0	+ 4,9	+ 0,0	+ 0,0

* Nem lehetett pontosan megállapítani a változás okát. Finnország fő okként koncepcionális változtatásokat jelölt meg.

Az említett hatásokat a tábla két csoportra bontva mutatja be.

A GDP egyes felhasználási tételei esetében látható, hogy az új módszerrel készült számítások a végső fogyasztásra kisebb hatást gyakoroltak, mint az állóeszköz-felhalmozásra és a külkereskedelmi adatokra. A változások alapvetően a koncepcionális, kisebb mértékben az alapadatok forrásainak változásából erednek.

A változást előidéző tényezők közül csupán néhányat emelünk ki:

- a beruházások értékébe beszámítanak az olyan szellemi javak, mint a számítógép-szoftverek beszerzése, az irodalmi és művészeti alkotások, a geológiai-kutatások és a – nem kizárólag katonai célra használható – katonai beruházások (honvédségi épületek, iskolák, kórházak, teherautók stb.) értéke;

- értékcsökkenést számolnak el olyan közösségi tulajdonban lévő infrastrukturális állóeszközök után is, melyeket korábban „örök életűnek” tekintettek (utak, gátak, hidak stb.);

- hatással volt a változatlan áras számításokra és a volumenindexekre az is, hogy az ármegfigyelésekben, illetve az árindek számításában gondosabban figyelték az új termékek és a minőségi változások kezelésére;

- az export- és az importadatokban az egyik legfontosabb változást a turizmus (a nem rezidens látogatók kiadásainak) elszámolása jelentette: a turisták belföldi vásárlása exportnak, míg a hazai turisták külföldi vásárlásai importnak minősülnek, s ez a változtatás jelentősen befolyásolja a külkereskedelmi adatokat (a múltban a magánfogyasztás az ország területén megfigyelt fogyasztást jelentette);

– a pénzügyi (bank, biztosítás stb.) tevékenységek számbavételénél lényeges változások következtek be az ESA95-ben, ezek hatását azonban még csak kismértékben tartalmazzák a jelenlegi adatok, mivel a módszertan részletes kidolgozása még nem fejeződött be.⁸

Az EU-tagországok többsége 1991–1997-re is visszavezette az új módszertan szerinti idősorát. Ennek eredményeit mutatja a 2. tábla.

2. tábla

Az ESA95 és az ESA79 folyó áras adatai közötti különbség százalékban

Tertület	1991.	1992.	1993.	1994.	1995.	1996.	1997.
	évben						
Eurózóna	:	:	:	:	+ 1,9	+ 1,7	+ 1,7
EU15	:	:	:	:	+ 2,0	+ 1,9	:
Belgium	+ 0,4	+ 0,5	+ 0,2	+ 0,2	+ 0,8	+ 0,3	+ 0,6
Dánia	+ 5,7	+ 5,8	+ 4,9	+ 6,7	+ 6,4	+ 7,4	:
Németország	+ 3,0	+ 2,5	+ 2,3	+ 2,0	+ 2,3	+ 1,8	+ 1,4
Görögország	:	:	:	:	:	:	:
Spanyolország	:	:	:	:	+ 4,4	+ 4,6	+ 4,8
Franciaország	+ 1,6	+ 1,8	+ 2,1	+ 1,5	+ 1,2	+ 1,0	+ 1,2
Írország	+ 3,1	+ 2,0	+ 1,7	+ 0,1	+ 0,2	+ 0,4	+ 1,8
Olaszország	+ 1,5	+ 1,6	+ 1,5	+ 1,3	+ 0,9	+ 1,2	+ 1,2
Luxembourg	:	:	:	:	:	:	:
Hollandia	:	:	:	:	+ 4,1	+ 3,9	+ 4,1
Ausztria	:	:	:	:	:	:	:
Portugália	:	:	:	:	+ 1,9	:	:
Finnország	+ 1,5	+ 2,2	+ 2,2	+ 2,0	+ 2,1	+ 2,4	+ 1,3
Svédország	:	:	+ 3,2	+ 3,7	+ 3,4	+ 3,8	+ 3,8
Egyesült Királyság	+ 1,8	+ 1,7	+ 1,6	+ 1,7	+ 1,6	+ 1,7	+ 1,6

Az eddig rendelkezésre álló adatok alapján a legnagyobb növekedés Dánia idősorában mutatkozik (4,9-7,4%), míg bizonyos években legkevésbé Írország és Belgium adatai módosultak (0,1-0,2%). Az is a módszertani váltás jelentőségét bizonyítja, mennyire különbözőképpen érintette az idősor egyes elemeit a felülvizsgálat.

MÓDSZERTANI VÁLTÁSOK A MAGYAR NEMZETI SZÁMLÁKBAN

Bár a magyar hivatalos statisztika még nem képes érvényesíteni az ESA95 valamennyi előírását, az 1990-es években több módszertani felülvizsgálatot hajtott végre. 1991-1992-ben érvényesítette az 1993-as SNA számos olyan ajánlását, amelyek alapvetők voltak a piaccgazdaságra való átállás következtében. 1997-ben 1995-ig visszavezetve további módszertani korrekciókra került sor. A kilencvenes évek elején végrehajtott változások törést jelentenek az idősorban, jórészt összehasonlíthatatlanná tették a korábbi idősorokat. Így ebben az esetben nem a szokásos statisztikai felülvizsgálatról beszélhetünk,

⁸ A FISIM, vagyis a pénzközvetítői szolgáltatás fel nem osztott díjának módszertani megoldásáról és alkalmazásáról jelenleg még élénk vita folyik a nemzeti számlák szakértői munkacsoport ülésén. Az új módszer alkalmazása a tagállamok jelentős részében lényeges hatást fog gyakorolni a GDP szintjére. A próbaszámítások szerint többnyire emelni fogja a GDP kimutatott szintjét.

hiszen a törés nem a statisztikai módszertanban, hanem az általa leírt gazdasági jelenségekben következett be. Gondoljuk meg, hogyan lehetne piaci áron értékelni a KGST keretében kialakult külkereskedelmi kapcsolatokat. Az 1990-es évek második felében átvezetett módosítások viszont alig befolyásolták a GDP szintjét, csupán annak belső szerkezetére gyakoroltak hatást, elsősorban a felhasználási oldalt, különösen a külkereskedelmi egyenleget és a készletváltozást érintették.

Az 1990-es évek közepére időzített korrekciókat a következő okok indokolták:

- az ESA 95 ajánlásaival való összhang erősítése;
- a szervezeti egységek, jogi helyzetében bekövetkezett változások nyomán követése (például a gazdálkodási forma változása);
- az üzleti számviteli változások figyelembevétele;
- a statisztikai adatgyűjtési rendszer módosulása.

A forrás oldalon a következő változások történtek.

1. Teljes körű felülvizsgálatra került az adók és támogatások statisztikai elszámolása. Ennek eredményeként az egyéb termelési adók csoportjából a termékadók közé került a szeszadó, a környezetvédelmi termékdíj, valamint a helyi adók közül az idegenforgalmi és a helyi iparüzési adó; az egyéb termelési támogatások csoportjából a terméktámogatások közé került az egyedi termelési árkiegészítés, az agrárpiaci és az agrártermelés költségeit csökkentő támogatás. A termékadók és -támogatások újracsoportosítása a nemzetgazdaság alapárak kibocsátását és hozzáadott értékét 37 milliárd forinttal csökkentette.

2. Ezzel párhuzamosan, a hosszú idősoros adatok elemzése alapján korrigálták a kisvállalkozások termelését és jövedelmét. Ez mintegy 50 milliárd forintnyi növekedést eredményezett a hozzáadott értékben.

3. A termékadók és a terméktámogatások egyenlege az átcsoportosítások, valamint a rádió és a televízió vállalati szektorba történt átsorolása miatt 12 milliárd forinttal nőtt. Ez a korrekció a piaci beszerzési áras hozzáadott értéket mintegy 25 milliárd forinttal növelte.

4. A termelési oldalon a szerencsejáték-bevételek elszámolásának módosítása, illetve a gépkocsihasználattól adódó és az egyéb munkavállalói jövedelemtételre folyó termelőfelhasználásból való kiszűrése a hozzáadott értéket 17 milliárd forinttal emelte. Erre a dezagregált adatok összehasonlíthatóságának biztosítása érdekében van szükség.

5. A korábban az államháztartási szektorba sorolt rádió és televízió, valamint a volt közúti igazgatóságok szervezeti átalakulása és az elszámolási technika ebből adódó változása csökkentette az államháztartási szektorban előállított hozzáadott értéket. Az előbbieket részvénytársasággá, az utóbbiakat közhasznú társaságokká alakultak át. A magyar rádió és televízió nemcsak az államháztartás szektorából került át a nem pénzügyi vállalatok szektorába, de a költségszintű értékelés helyett a vállalati szektor általános szabálya szerinti elszámolási szabály vonatkozik rájuk. Ennek következtében az államháztartási szektor hozzáadott értéke 13 milliárd forinttal csökkent, míg a vállalati szektoré 22 milliárd forinttal nőtt.

6. Jelentősen csökkent a háztartásokat segítő nonprofit intézmények kibocsátása annak következtében, hogy a szektorból leválasztották a vállalatokat segítő, valamint az államháztartás által irányított és döntően állami forrásból finanszírozott intézményeket. Ezeket a vállalati, illetve az államháztartási szektorba sorolták át.

A felhasználási oldalt érintő változások a következők voltak.

7. Lényegesen változtak a külkereskedelmi tevékenység export–import adatai, döntően annak hatására, hogy 1995-től sikerült az ipari vámszabad területek forgalmáról olyan adatokat gyűjteni, amelyek ezek külkereskedelmét a rezidens termelőkével azonos módon veszik számba. A szokásos vámstatisztikai megfigyelés ugyanis a vámszabad területeket nem tekinti a nemzetgazdaság részének. A változtatás hatására a külkereskedelmi egyenleg 50 milliárd forinttal javult.

8. 1995 óta a beruházásstatisztikai adatgyűjtés lehetővé teszi olyan tételek – immateriális javak beszerzése, eszközök apportja és a pénzügyi lízing, használt eszközök forgalma – beépítését a rendszerbe, amelyekre ko

rábban nem volt lehetőség részben azért, mert a számviteli elszámolások ezeket nem tekintették az állóeszköz-felhalmozás részének. A korábban közzétett értékhez képest a változtatások +65 milliárd forint eltérést jelentettek.

9. A közösségi fogyasztás értéke a rádió és televízió átsorolása – és egy téves adatközlés javítása miatt – 12 milliárd forinttal csökkent.

10. Az új módszertan szerint a gazdálkodó szervezetek tulajdonában lévő gépkocsik magáncélú felhasználását nem folyó termelőfelhasználásnak, hanem lakossági fogyasztásnak tekintik. Nagyrészt ez okozta a lakossági fogyasztás 9 milliárd forintos emelkedését.

A korrekciók miatti változásokat a 3. tábla szemlélteti.

3. tábla

Az 1995. évi GDP módosulása a módszertani változások következtében

Megnevezés	Korábban közölt adat	Korrigált adat	Eltérés	
	millió forint		százalék	
	Termelési oldal			
Vállalatok	2 529 344	2 586 182	56 838	+2,2
Pénzügyi vállalkozások	260 283	259 002	-1 281	-0,5
Államháztartás	948 463	935 333	-13 130	-1,4
Háztartások	1 110 166	1 113 078	2 912	+0,3
Háztartásokat segítő nonprofit intézmények	44 286	39 260	-5 026	-11,3
Ágazatok összesen	4 892 542	4 932 855	40 313	+0,8
FISIM	-219 248	-219 248	0	0,0
GDP (alapáron)	4 673 294	4 713 607	40 313	+0,9
Termékdadók egyenlege	888 571	900 435	11 864	+1,3
<i>GDP (piaci áron)</i>	<i>5 561 865</i>	<i>5 614 042</i>	<i>52 177</i>	<i>+0,9</i>
	Felhasználási oldal			
Háztartások fogyasztása	3 714 957	3 723 955	8 998	+0,2
Közösségi fogyasztás	629 472	617 700	-11 772	-1,9
Összes végső fogyasztás	4 344 429	4 341 655	-2 774	+0,1
Állóeszköz-felhalmozás	1 059 559	1 125 389	65 830	+6,2
Készletváltozás és egyéb nem specifikált felhasználás	279 688	218 346	-61 342	-21,9
Összes felhalmozás	1 339 247	1 343 735	4 488	+0,3
<i>Belföldi felhasználás</i>	<i>5 683 676</i>	<i>5 685 390</i>	<i>1 714</i>	<i>+0,03</i>
Export	1 914 813	2 091 797	176 984	+9,2
Import	2 036 624	2 163 145	126 521	+6,2
Külkereskedelmi egyenleg	-121 811	-71 348	50 463	+41,4
<i>Bruttó hazai termék (GDP) összesen</i>	<i>5 561 865</i>	<i>5 614 042</i>	<i>52 177</i>	<i>+0,9</i>

A tábla adatai szerint, az 1995-ös adatokon végzett korrekciók mintegy 0,9 százalékkal növelték a GDP értékét. A változtatások szinte kivétel nélkül módszertani jellegűek voltak, kisebb mértékben technikai korrekcióknak tekinthetők, míg új információk, adatforrások beépítésére és adatkijavításokra valójában nem került sor.

Összefoglalva, a tájékoztatási stratégia szempontjából a statisztikai adatok felülvizsgálatának két ismérve számít:

- az, hogy rendszeresen számíthatunk-e a felülvizsgálatra, vagy arra csak kivételes esetben kerül sor;
- lehet-e az adatközlést megelőzően bizonyosra venni, hogy változni fog az adatok értéke, vagy a változás ténye előre nem látható.

Az új, pontosabb adatokra rendszeresen számíthatunk, és szinte bizonyosra vehető, hogy ezek hatására meg kell változtatni a már közölt adatokat. A jelentős módszertani változások rendkívüli jellegűek, és jó előre tervezhetők. Az adathibák kivételesen fordulnak elő, és a hiba mértéke csak utólag derül ki. Az adatértelmezési, technikai korrekciókra okot adó eseményekre rendszeresen sor kerülhet, de nem minden esetben kell azok hatását az adatokon átvezetni. A tanulmányban foglaltak azt is érzékeltetni kívánták, hogy a magyar statisztikai szolgálatban, az elkövetkező időszakban szinte biztosan, fel kell készülni nemcsak arra, hogy a szokásosnál gyakrabban kell az adatokat felülvizsgálni, hanem arra is, hogy követve a statisztikai rendszer változását, várhatóan maga a felülvizsgálati politika is többszöri felülvizsgálatra fog szorulni.

IRODALOM- ÉS FORRÁSJEGYZÉK

- DEN BAKKER, G. P. – DE GIJT, J. – VAN ROOIJEN, R. A. M. (1996): *New revision policies for the Dutch National Accounts. Economic and Social History in the Netherlands, Historical National Accounts in the Netherlands, The Netherlands Economic History Archives, Amsterdam.*
- BEADLE, J. (1999): National Accounts revision to data and revisions policy. (Kézirat.)
- BEDÉKOVICS I. (1994): A háztartási szektor elszámolásának módszertana. *Statisztikai Szemle*, 72. évf. 7. sz. 530–546. old.
- ESA 1995 – Transmission programme of data (1997).* EUROSTAT, Luxembourg.
- European system of accounts – ESA1995 (1996).* EUROSTAT, Luxembourg.
- FAZEKASNÉ KOVÁCS K. (1994): A vállalati szektor jövedelemeloszlása. *Statisztikai Szemle*, 72. évf. 8–9. sz. 613–624. old.
- Handbook on Quarterly national accounts (1999)* EUROSTAT, Luxembourg.
- DR. HÜTTL A. (1994): A magyar nemzeti számlák új adatforrásai és módszerei. *Statisztikai Szemle*, 72. évf. 6. sz. 469–480. old.
- Implementation of ESA95 Major progress by statisticians to improve the National Accounts (Eurostat New Release, No 69/99, 16 July 1999). További információk az EUROSTAT 1999. júniusi „Meeting of the Working Party on National Accounts” ülésére benyújtott „Evaluation of the first transmission of ESA95 data” illetve az „Impact of the transition from ESA79 to ESA95” jelentéseiben állnak rendelkezésre.
- Magyarország nemzeti számlái (1994).* Adatforrások, módszerek és számítások. KSH–OECD, Budapest.
- Magyarország nemzeti számlái (1998).* Adatforrások, módszerek és számítások – Felülvizsgált kiadás. KSH–OECD, Budapest.
- DR. POZSONYI P. (1994): A termelési számla. *Statisztikai Szemle*, 72. évf. 10. sz. 744–760. old.
- Revision policy of series of National Accounts (2001).* EUROSTAT, (Munkaanyag.)
- WROE, D. (1993): Handling revisions in the National Accounts. *Economic Trends*, 480 sz.

SUMMARY

The paper provides a broad overview on the major issues, which should be considered when formulating the revision policy of an official statistical service. Four reasons for the data revision are distinguished: new data sources, methodological changes, late discover of errors in the data and other technical corrections for example due to the re-estimation of parameters of the time series estimated by models. Different treatments are appropriate in the case of short-term business cycle indicators, annual basic statistics and national accounts. The main differences and the basic ideas behind them are analysed. Both regular and exceptional revisions must be planned and the users should always be informed in advance about the fact and the causes of the revisions. The paper illustrates the implementation of the revision policy using examples from the practice of the Hungarian statistical service.

AZ EMBERI ÉLET ÉRTÉKE MAGYARORSZÁGON

ADORJÁN RICHÁRD

A tanulmány egy sokszor és sok irányból támadott kísérletre vállalkozik: megmérni az emberi élet értékét. Azokban az országokban, ahol a közpolitika alakításában már évtizedekkel ezelőtt gyökeret vertek olyan modern technikák, mint amilyen a költség–haszon elemzés, nem idegenkednek annyira – sokszor még a közgazdászok sem – az effajta számításoktól. Nem célunk ugyanis mérlegre állítani az egyes emberek életét, hanem a lineáris regressziószámításra épülő ún. hedonikus ármódszert alkalmazva egy speciálisan definiált statisztikai emberi élet értékét határozzuk meg.

A felállított modelleken elvégzett számítások azt jelzik, hogy ez az érték 1998-ban Magyarországon a 78-393 millió forintos intervallumban volt, legmegbízhatóbbnak a 250 millió forint körüli érték tekinthető. Ez hozzávetőleg fele az Egyesült Államokban és Nagy-Britanniában – más kultúrával, jövedelmi és vagyoni viszonyokkal, s más preferenciaterkép-pel rendelkező populáción – hasonló eljárással nyert eredményeknek.

TÁRGYSZÓ: Költség–haszon elemzés. Közpolitika-elemzés. Hedonikus ármódszer.

Az emberi élet értékének számszerűsítése nem tartozik a közgazdaságtan klasszikus feladatai közé, s nem is nagyon népszerű foglalatosság. Joggal, főként abból a szemszögből, ahogyan *prima facie* viszonyulunk e kérdéshez: az emberek élete valóban nem lehet csere tárgya, áru, még csak nem is jószág, értéke nem gazdasági kategória.

Az állam pénzt áldoz polgárainak védelmére, ugyanakkor lemond más javak juttatásáról, például nem épít iskolát, vagy nem emeli az egészségügyi dolgozók fizetését. Minek alapján dönthetünk azonban arról, hogy az igencsak szűkre szabott büdzséből a számtalan lehetséges szükséglet közül végül is mire és mennyit költsünk? E kérdés megválaszolásában segíthetnek a költség–haszon elemzés technikái, eljárásai és eredményei, például az emberi élet értékének meghatározása.¹

A KIINDULÓ PROBLÉMA

A közgazdaságtan, az ún. intézményi közgazdaságtani irányzattal, a hetvenes években kezdett behatolni az élet olyan szféráiba, amelyek sokak szerint nem tartoznak s nem is

¹ A kutatás, melyből e tanulmány íródott, részben a United States Information Agency által finanszírozott SSCD (Social Sciences Curriculum Development Program for Selected Central European Universities) program anyagi támogatásával valósult meg. Természetesen az itt kifejtett gondolatokért e szervezetek nem vállalnak felelősséget.

tartozhatnak e diszciplína vizsgálati körébe. „Az új irányzat képviselői szerint az ember nem csupán a szűk értelemben vett gazdasági döntési helyzetekkel egyénileg szembenézve igyekszik törekvéseit értelmesen, racionálisan érvényesíteni, hanem családi kapcsolataiban, politikai aktivitásában, netán még akkor is, amikor a szegények vagy a mozgássérültek javára áldoz, irodalmi divathóbortot kap fel, vagy érzelmeit igyekszik kiélni.” (Bara–Szabó; 1997. 45. old.)

Jelen kutatásunkat, megközelítési módszerünket és elemzési technikánkat mégsem sorolnánk e kétségkívül vonzó és kihívó irányzathoz, még ha a vizsgálat tárgya (a homo sapiens, vagyis maga az ember) ezt lehetővé tenné. Érdemesnek tűnik tehát pontosan körülhatárolni a problémát, s a megtalált megoldási módot, amellyel a továbbiakban foglalkozunk.²

Egy heterogén társadalom sokféle igénnyel lép fel az állammal szemben, amely igények bizonyos, jól vagy kevésbé jól működő mechanizmusok révén konkrét programokká, projektekévé aggregálódnak. Ezen programok megvalósításának egyetlen közgazdasági kritériuma van: mégpedig az, hogy a nettó társadalmi hasznuk (vagyis a társadalmi haszon és társadalmi költség különbsége) pozitív, s az alternatív programokénál nagyobb legyen.

A lehetséges, relatív előnyöket és hátrányokat felmutató módszerek közül a költség–haszon elemzés bizonyult a legélethez közelebbinek.³ Ennek lényeges tulajdonsága, hogy egy adott gazdasági rendszer értékelése során mind a hasznoknak, mind a költségeknek pénzértékben való kifejezését követeli meg. Ez azonban sok esetben lehetetlen feladatnak látszik. Az emberi élet értékét⁴ (is) éppen azért kell pénzértékben meghatározni, hogy ezáltal összevethetővé váljék egyéb tényezők értékével. Látnunk kell, hogy az állam különböző szerveinek döntéseiben létezik implicit módon átváltás az emberi élet és egyéb javak, például a művelődés, a kulturáltság értéke között, hiszen ellenkező esetben mindaddig nem épülhetne egyetlen könyvtár sem, amíg minden úttest mellé korlátot nem szereltünk, minden folyóhoz gátat nem építettünk, s egyáltalán, minden lehetséges eszközt fel nem használtunk az emberi élet védelme érdekében. Ezek a döntések azonban, éppen tudományos megalapozatlanságuk miatt, jobbra esetlegesek, ki vannak téve a politikai szereplők saját preferenciáinak, melyek csak ritkán felelnek meg a választók hasznosságrendszerének.⁵

² Ezúton szeretnék köszönetet mondani mindazoknak, akik segítséget nyújtottak a megfelelő irodalom kiválasztásában, illetőleg annak elérésében. Külön köszönettel tartozom *David H. Greenbergnek*, *Sherwin Rosennek*, *Pauline M. Ippolitonak*, *Shelby Gerkingnek* és *Ann Fishernek*, akik olyan tanulmányokkal láttak el, amikhez nélkülük lehetetlen lett volna a hozzáférés.

³ Mind magáról a költség–haszon elemzésről, mind a lehetséges egyéb eljárásokról kitűnő áttekintést nyújt *Boardman–Greenberg–Vining–Weimer* (1996).

⁴ Az emberi élet „közgazdasági értékének” meghatározására először *Dublin és Lotka* (1930) tett kísérletet. Véleményük szerint az tekinthető adekvát eljárásnak, ha meghatározzuk az adott egyén jövőbeli nettó jövedelmeinek jelenértékét. *Reynolds* (1956) kísérlete meg először bizonyos „megfoghatatlan” tényezők számszerűsítését, elsősorban a halálessel járó közúti balesetekkel kapcsolatosan. *Abraham és Thedié* két kategóriát is felállított az ilyen baleseteknél elszenvedett károk elkülönítésére: a „közgazdasági komponenseket” és a „szubjektív komponenseket”. Ez utóbbi csoportban olyan tényezők szerepeltek, mint 1. a rokonok szenvedése, 2. a társadalom szubjektív költségei, 3. az életben maradt áldozatok szenvedése, 4. az életben maradt áldozatoknak az élet élvezetében bekövetkezett csökkenése, s 5. az emberek afelőli félelme, hogy a jövőben hasonló balesetek áldozatai lehetnek. *J. Dreze* volt az első, aki elődeitől egészen eltérően közelítette meg a kérdést, a ma már széles konszenzusra épülő módszertan elméleti alapjait rakva le ezzel. Elemzését két lépésben hajtotta végre: először az egyén szintjén vizsgálta meg a biztonság javításának értékét, majd ezen értékelési problémát kollektív döntéshozatali kontextusba helyezte. Így okoskodott: ha az egyének által olyan situációkban meghozott döntések, amelyeknél bizonytalanság van jelen, kielégítik a szabályosság és konzisztencia követelményét, akkor a közgazdasági eszköztár tökéletesen alkalmas arra, hogy megpróbálja megállapítani, hogy az emberek mennyit hajlandóak áldozni a biztonságukban bekövetkezett bizonyos növekedésért cserébe. Ezen elméletekről, s az itt meg nem említettekről is részletes történeti áttekintést nyújt nagyhatású könyvében *Jones–Lee* (1976).

⁵ Vö. *Downs* (1990) ismert elméletével s a nyomában járó szakirodalommal.

Miről is van szó? Azt kívánjuk meghatározni, hogy a társadalom (vagy egy adott populáció) tagjai hogyan értékeli saját életüket, így ennek preferenciarendszeréből kell kiindulnunk. Mára abban is egyetértés mutatkozik a közgazdász-társadalomban (*Viscusi*; 1986. 207. old.), hogy az emberi életen ilyen kontextusban csakis a statisztikai emberi életet érthetjük, ennek értékét pedig az emberek mindenkor fizetési hajlandóságára építve határozhatjuk meg. Nevezetesen a vizsgált populáció körében összegeznünk kell, hogy az egyének mennyi pénzt hajlandók feláldozni a halál kockázatának bizonyos csökkentése érdekében, vagy mennyi pénzt hajlandók elfogadni kompenzációként e kockázat növekedésének fejében.

A gondolati ívet e ponton érdemesnek tartjuk megszakítani egy technikai jellegű részszel, a használt kulcsfogalmaink operacionalizálásával, helyesebben, szándékaink szerint a gondolatmenetet éppen ezzel tesszük folyamatossá.

Kockázat

Az emberi életet veszélyeztető mindennemű kockázatot az élet elvesztésének (vagyis a halál bekövetkeztének) matematikai valószínűségeként definiálunk.

Azonosítható emberi élet

Az azonosítható ember egy létező vagy létezett, a jelenben jól beazonosítható személy, mint például e sorok írója. Az ilyen emberi élet értékének meghatározása nem célunk, s az általunk követett filozófiai megfontolások és közgazdasági logika erre nem is ad lehetőséget. Semmiképpen sincsen tehát arról szó, hogy egy azonosítható ember életének értékével bármilyen módon foglalkozunk.

Statisztikai emberi élet

A statisztikai ember fiktív ember, nem létezik, nem létezett, s nem is fog létezni (szemben az anonim emberrel!), mindössze egy valószínűségekből álló összeg. A statisztikai emberi élet mindig csak adott populációban, embercsoportban értelmezhető. Fogalma együtt jár a halál jelenségével, mégpedig (egy adott időintervallumon belül) a populáció egyes tagjaira egymástól függetlenül leselkedő halál bekövetkezési valószínűségeinek 1-et adó összege éppen egy statisztikai emberi élet.

Világosabbá válik mindez egy példán keresztül. Ha egy védőberendezés egy ezer fős populáció minden tagjánál egy adott évben 2 ezrelékről 1 ezrelékre (vagyis 1 ezrelékponttal) csökkenti a halál valószínűségét, akkor azt mondjuk, hogy a védőberendezés 1 statisztikai emberi életet mentett meg az adott évben (hiszen $1000 \times (1/1000) = 1$).

Statisztikai emberi élet értéke

A statisztikai emberi élettel önmagában sohasem dolgozunk, sokkal lényegesebb számunkra a statisztikai ember életének értéke. Ez azzal a pénzösszeggel egyenlő, amit az életüket veszélyeztető kockázatnak kitett tagok összességében hajlandók fizetni azért, hogy a halál – az egyes egyéneknél függetlenül tekintett – bekövetkezési valószínűségének külön-külön kismértékű, együttesen 1 összegű csökkenése következzen be.

Az előző példánál maradva: ha az egyszerűség kedvéért mind az ezer ember 10 000 forintot hajlandó fizetni az adott védőberendezésért (vagyis a kockázat egy ezrelékponttal való csökkenéséért), akkor a statisztikai emberi élet értéke $1000 \times 10\,000 = 10\,000\,000$ forint az adott populációban.

Anonim emberi élet

A statisztikai emberi élettel szemben az anonim emberi élet nem fiktív élet, létezett és létező személy, csupán jelen pillanatban senki nem ismeri kilétét, vagyis személye – kizárólag az időtényező miatt – azonosíthatatlan, a jövőben azonban azonosíthatóvá válik. Tulajdonképpen matematikai értelemben egy valószínűségi változóról van szó, amely jellegéből fakadóan a priori nem, csupán a posteriori kap konkrét értéket.

Példánknál maradva egy védőberendezés akkor ment meg egy anonim emberi életet, ha anélkül pontosan egy ember vesztene el életét egy adott évben, akinek kilétéről azonban még nem tudunk semmit mondani.

Anonim emberi élet értéke

Az anonim emberi élet értéke hasonlóan adódik, mint a statisztikai emberi élet értéke, azzal a különbséggel, hogy a populáció tagjainak most a pontosan egy megmentett (de kilétében előre nem ismert) életet kell értékelniük, vagyis azon pénzösszegeket kell összeadnunk, amit a populáció tagjai külön-külön fizetnének azért, hogy elkerüljék a mindenkire egyenlően kis valószínűséggel leselkedő, de biztosan bekövetkező egy halálesetet.

Fontos hangsúlyozni, hogy e helyütt kizárólag a statisztikai emberi élet értékének vizsgálatával foglalkozunk. A fogalom definíciójából adódik, de érdemes kiemelni, hogy számunkra csakis az élet elvesztésének kockázata, ennek értékelése számít. Az értékelendő élet ilyenformán egy van–nincs jellegű kétértékű változó, semmilyen módon nem jelennek meg számításainkban az élet minőségére utaló paraméterek. Nem azért, mert ezek nem fontosak,⁶ hanem azért, mert a közgazdasági elméletben megegyezés útján kialakult statisztikai emberiélet-érték fogalom az ilyen változókat nem tudja és nem is akarja beépíteni gondolat- és számítási rendszerébe.

Nem állítható ugyanakkor, hogy ez az egyetértés megtámadhatatlan, ellenkezőleg: éppen a kérdéskör érzékenységénél fogva, könnyen támadható s gyakorta támadott. Az első sorban *Schelling* (1968) és *Mishan* (1971) által megalapozott elméletet, s az ezen alapuló eljárást *Acton* (1976) az aktuális jövedelem- és vagyoneeloszlás méltánytalanságára hivatkozva utasítja el, mivel az eljárásban – egyszerűen fogalmazva – a nagyobb vagyonnal és jövedelemmel rendelkezők szava, „szavazata” többet nyom a latba, hiszen a fizetési hajlandóság alapját a fizetési képesség jelenti. *Acton* ezért a fizetési hajlandóságon alapuló metódusokat nem tartja elfogadható alapnak a javak előállításának és elosztásának meghatározásához.

Talán a legzajosabb vitát kiváltó kritikát *J. Broome* (1978) intézte *Mishanék* gondolatmenete ellen, melynek lényege, hogy mivel ex ante csupán annyit tudunk, hogy várhatóan hány áldozatot követel majd a projekt, azt azonban nem, hogy pontosan hányan és kik lesznek azok, ezért a program költsége a halál kockázatában bekövetkezett növekmény (s így ezért kell kárpótolni az embereket), nem pedig maga a halál, ennek értékét pedig már tudjuk mérni, s mérhetjük is. Broome szerint a döntés szempontjából nem lehet releváns kérdés az, hogy mikor derül ki, hogy kinek kell meghalnia. Tudjuk, vagy nem tudjuk, hogy ki lesz az: mindegy, a lényeg, hogy meghal-e valaki, illetve hányan halnak meg. Érvelése alátámasztására, többek között, az alábbi abszurd példát hozza: azok nevét, akik meg fognak halni egy projekt miatt, egy lezárt doboz tartalmazza. Ha a dobozt felnyitjuk, a projektet – *Mishan* szerint – azonnal el kell vetnünk, a felnyitás előtt azonban még elfogadhatjuk.

Amikor tehát, a bevett módszernek megfelelően, a statisztikai emberi élet értékét akarjuk meghatározni, Broome szerint helytelenül járunk el. Az élet értékét ugyanis az anonim emberélet mérésével kellene meghatároznunk (ekkor biztos ugyanis az élet elvesztése), ami viszont elméletileg lenne helytelen, hiszen az ex ante tudott, bár személyében anonim halál utáni világállapotot (s ezzel együtt tehát a halál miatti értékvesztéset, így az élet értékét) a társadalom axiómaszerűen nem értékelheti különbözően ugyanezen halál utáni ex post (tehát már nem anonim!) állapot ex ante értékelésénél.⁷ Ez utóbbi értékelés azonban, mint azt már korábban beláttuk, elméletileg lehetetlen.

Jelen dolgozatunknak semmiképpen sem célja, hogy állást foglaljunk ezekben a vitákban, illetőleg annyiban tesszük ezt meg, amennyiben elfogadjuk a kialakult konszenzus módszertani ajánlását, s ennek megfelelően végezzük el számításainkat. Azt valljuk ugyanis, hogy egy még oly bizonytalan és támadható elvekre épülő módszer alapján végzett számítá-

⁶ Számos közgazdász műhely foglalkozik is olyan jellegű számításokkal, amelyeknek középpontjában éppen az életminőség értékelése áll.

⁷ Ez természetesen már attól függ, hogy milyen axiómákat fogadunk el a társadalmi értékelésre, vagyis a társadalmi jóléti függvényre vonatkozóan. A kérdés részletes vizsgálatát lásd: *Ulph*; 1982.

sok is értékes és használandó eredményhez juttatják a közpénz hatékony elköltésén fáradozó döntéshozókat.⁸ Kutatásunknak pedig éppen e gyakorlati használhatóság a célja.

Álljon itt egy közérdeklődésre számot tartó magyar probléma példája. Magyarország fővárosának, s vitathatatlan közlekedési centrumának gépjárműforgalmi tehermentesítése régóta jogos igény mind a szakemberek, mind a lakosság körében. Az 1990-es évek derekán átadott M0-ás körgyűrű a Budapestre sugárszerűen beérkező autópályák és más főutak közötti összeköttetést volt hivatott létrehozni a város elkerülésével. Valószínűleg több kedvezőtlen körülmény (konstrukciós hibák, forgalom-félrebecslések, járművezetői szabálytalanságok stb.) szerencsétlen összjátékának következményeként a néhány tíz kilométeres útszakaszon ritkán tapasztalható gyakoriságot értek el a baleseti s ezáltal halálozási statisztikák. A tényleges okok szempontunkból nem lényegesek, csupán az a tény, hogy meghatározott költségek árán bizonyosan csökkenteni lehetne a várható, vagyis a statisztikai halálesetek számát, másként fogalmazva statisztikai emberi életeket lehetne megmenteni. E költségek az alternatív megoldásoktól függően, s jellegüket tekintve is többfélék: átépítési költségek, szabályozási költségek, a lassabb vezetés miatti időkiesésből eredő költségek stb.

Mármost a kérdés úgy merül fel, hogy y kiszámítható nagyságú költségek mellett érdemes-e megmenteni x statisztikai emberi életet (vagyis a halál bekövetkeztének valószínűségét ennek megfelelően csökkenteni). Erre a kérdésre a döntéshozóknak kell megadni a választ. Jelen számításaink eredménye éppen abban segít, hogy a döntéshozók a társadalmi preferenciákra támaszkodva össze tudják hasonlítani a kiszámított költségeket a statisztikai emberi életek megmentése által nyert hasznokkal. Egy ilyen összevetést nyilvánvalóan csak akkor lehet (helyesebben akkor érdemes) megtenni, ha a két kategória mértékegysége azonos, ami jelen esetben megegyezik a pénzértékkel, vagyis magyar forint.

Fontos látnunk, hogy a döntéshozó csak támpontot kap döntéséhez. Sokféle ok lehet arra, hogy eltérjen a rideg számok nyújtotta ítélettől: más szubjektív súlyokat helyez az egyes költség- vagy haszontételekre, politikai tényezők módosítják a döntést, közrejátszhat a média vagy a közvélemény nyomása vagy éppen érdektelensége stb. Ezek a tényezők azonban már kívül esnek azon a mezőn, ahol mi mozogni kívánunk. Mihelyt meghatároztuk a statisztikai emberi élet értékét, bennünket tovább már „nem érdekel” a probléma sorsa. Számításaink csupán egy, bár igen fontos, láncszemet biztosítanak a teljes láncolathoz, e nélkül azonban – hisszük – használhatatlanná válik a lánc és esetlegessé a döntéshozatal.

A STATISZTIKAI EMBERI ÉLET ÉRTÉKÉNEK MÉRÉSI MÓDSZEREI

Az elmúlt évtizedek során a statisztikai emberiélet-érték empirikus meghatározásának három ága terjedt el: a munkaerő-piaci megközelítés,⁹ az egyéb piaci megközelítés, s az ún. contingent valuation, azaz a feltételes értékelés módszere. A következőkben ezeket ismertetjük röviden.

⁸ Gondoljunk bele, hol tartana a közgazdaságtan fejlődése, s az erre épülő mikro- és makrogazdaság alakulása, ha csak a „megtámadhatatlan” elvekre és módszerekre építkezett volna néhány száz éves történelme során e diszciplína.

⁹ A munkaerő-piaci megközelítésre, másképpen a kockázatért járó pótlólagos munkabéren alapuló tanulmányok (angolul ez sokkal szebben hangzik: „compensating wage differential studies”) módszerére használatos még a munkabérkockázat („wage-risk”), illetőleg a hedonikus ármódszer („hedonic price method”) kifejezés is. A továbbiakban ezeket szinonimákként kezeljük.

A hedonikus ármódszer

A modern közgazdaságtan skót atyjának, *A. Smith*-nek ismert művében, A nemzetek gazdagságában jelenik meg talán elsőként az a gondolat, hogy sajátos körülmények „...egyes foglalkoztatási területeken kárpótlást jelentenek azért, hogy a pénzbeni nyereség csak kisebb, míg másokon ellensúlyozzák az azokon elérhető nagyobb nyereségben rejlő csábítást...” (*Smith*; 1992. 109. old.). E jelenség az alapja a kockázattal járó pótlólagos munkabéren alapuló számításoknak.

A munkaerőpiac közgazdasági alapelmélete értelmében a mindenkori munkabérek szintje olyan, hogy amellet éppensúlyban van az aggregált piaci kereslet (vagyis a munkaadók kereslete) és a piaci kínálat (vagyis a munkavállalók kínálata). Az is világos továbbá, hogy az adott munkabér ellenében betöltött állás nem más, mint tulajdonságok halmaza. A munkaadó e tulajdonságok függvényében hajlandó bizonyos pénzösszeget lemondani az állás betöltője számára csakúgy, mint a munkavállaló is e jellemzői szerint alakítja ki véleményét az általa megkívánt kompenzációról. Minden foglalkozás, illetve állás alapvető sajátosságai közé tartozik annak biztonságos volta, azaz, hogy milyen kockázattal jár a munka elvégzése az emberi életre. Azokban az állásokban, amelyeknél *ceteris paribus*, vagyis egyéb körülményeiben megegyező állásokhoz képest, nagyobb a kockázat, nagyobb kell lennie a fizetésnek is, tehát a munkabérnek tartalmaznia kell egy kockázati prémiumot. Ez elsősorban két ok miatt van így.

- A nagyobb kockázattal járó állást csak magasabb fizetés ellenében hajlandó betölteni a munkavállaló.
- A nagyobb kockázat elkerülése költséggel jár a munkaadónak, ezt kompenzálандó, a nagyobb biztonság fejében csak kisebb fizetést hajlandó biztosítani.

A magasabb kockázattal járó állások tehát *ceteris paribus* magasabb kompenzációt eredményeznek, ám a valóságban ilyen jellegű összehasonlításokra ritkán nyílik mód, hiszen minden egyes vizsgált munkabér háttérében egy sajátos jellemzőkkel bíró egyén és egy ugyancsak egyéni jellegzetességeket viselő munkahely, állás húzódik meg. Feladatunk tehát annak meghatározása, hogy miként hatnak a munkabérnek, mint függő változóknak a nagyságára (w_i) a személyiségjellemzők ($x_{im}, m = 1, \dots, M$), a különböző egyéb, köztük az adott állással, kivéve annak kockázatával kapcsolatos tényezők ($z_{ik}, k = 1, \dots, K$), illetve kiemelve az adott állás, foglalkozás kockázata (p_i). Az alábbi egyenlet képezi tehát az ilyenkor alkalmazandó regressziószámítás alapját:

$$w_i = \alpha + \sum_{m=1}^M \Psi_m x_{im} + \sum_{k=1}^K \zeta_k z_{ik} + \gamma p_i + u_i,$$

ahol α konstans, u_i a véletlen hibatermék, Ψ_m , ζ_k és γ pedig az adott változókhoz tartozó koefficiensek, amelyek közül a kutatási kérdés szempontjából az utóbbi az érdekes, ez jelzi ugyanis, hogy mennyi pótlólagos jövedelmet „ér” a kockázat növekedése, illetőleg csökkenése a munkaerőpiacon. Előbbi esetben értelmezésünk szerint a dolgozó ekkora munkabér-növekedésért cserébe vállalja a pótlólagos kockázatot, utóbbi esetben ennyit hajlandó fizetéséből feladni ahhoz, hogy bizonyos mértékben csökkenjen a kockázata. Ha mármint ezen értékeket lineárisan extrapoláljuk, akkor kapjuk meg azt az értéket, amit a kollektíva egy statisztikai emberi életnek tulajdonít.

E megközelítés támogatói elsősorban azzal érvelnek, hogy tényleges, a piac próbájának kitétt adatok megfigyelésével kalkulálható a statisztikai emberi élet értéke. Mivel a

jelen tanulmány éppen egy ilyen számítás elvégzésére tett kísérletet mutat be, így a módszer további előnyeivel és hátrányaival a későbbiekben még foglalkozunk.

A másik két megközelítés¹⁰

Az egyéb kinyilvánított preferencia-vizsgálatra épülő kutatások közös jellemzője, hogy szintén valós (de nem munkaerő-) piaci döntések vizsgálata révén kívánják a kockázat–pénz tradeoff-ot meghatározni. A fő nehézség abban jelentkezik, hogy a kockázattal közvetlenül nem kereskednek a piacon, így, lehetséges megoldásként, olyan termékek piaca vizsgálható, amelyek valamilyen szoros kapcsolatban állnak a biztonsággal, illetve az életveszéllyel. R. Dardis (1980) kutatásában például a füstjelzőpiacot veszi górcső alá, azt vizsgálja tehát, hogy az emberek, közösségek mennyit hajlandóak ténylegesen áldozni egy olyan berendezésre, amely kimutatható mértékben növeli életük biztonságát.

A contingent valuation esetén egyfajta kérdőíves megközelítésről van szó, amíg azonban a „valódi” kérdőíves módszer lényege, hogy a kérdező nem közöl információt az interjúalannal, hanem azt akarja megtudni, hogy a kérdés pillanatában meglévő információi alapján hogyan reagál a kérdésekre, addig a contingent valuation eljárás sajátosan kombinálja ezt a döntési kísérlet módszerére jellemző vonásokkal, elsősorban a fordított irányba is végbemenő információáramlással (Csontos–Kornai–Tóth; 1996. 9. old.).

Arról van tehát szó, hogy a kérdezőbiztos bizonyos hipotetikus piaci vagy egyéb döntési helyzet elé állítja a megkérdezettet, akinek tehát esetünkben határoznia kell: mennyit lenne hajlandó fizetni (vagy mennyi kompenzációt lenne hajlandó elfogadni) egy adott (általában igen kis mértékű) fizikai kockázat csökkentéséért (növeléséért), például mennyivel nagyobb összeget szánna arra, hogy egy görögországi körutazást egy bizonyíthatóan biztonságosabb autóbusszal tegyen meg. A válaszokból becsülhető a keresett statisztikai emberiélet-érték.¹¹

ALAPHIPOTÉZIS

A kutatási hipotézisek megfogalmazása esetünkben nem kézenfekvő feladat, mert a kutatás jellegéből adódóan, az sokkal inkább mennyiségi összefüggéseknek, mint ezen összefüggések létének, vagy bizonyos állítások igazságtartalmának vizsgálatára irányul. Az empirikus kutatások általunk alkalmazott munkaerő-piaci megközelítésének alaphipotézise a következőképpen fogalmazható meg: a munkavállalók által végzett munka kockázatának mértéke befolyásolja az érte kapott munkabér nagyságát. E kapcsolat pozitív irányú: nagyobb kockázathoz nagyobb munkabér tartozik.

A hipotézis vizsgálatára azért van szükség, mert tudni akarjuk, hogy egyáltalán van-e lehetőség Magyarországon is a munkabérek elemzésének segítségével mérni a statisztikai emberi élet értékét, működnek-e az ehhez szükséges mechanizmusok. Ha a hipotézist el kell vetnünk, akkor e módszer Magyarországon nem alkalmazható. Amennyiben a hipoté

¹⁰ A módszerek részletesebb tárgyalását lásd például: Fisher–Chestnut–Violette (1989).

¹¹ A Függelék F.1. táblája bemutatja a legjelentősebb kutatások empirikus eredményeit. Látható, hogy ilyen irányú vizsgálatokat elsősorban az angolszász országokban, az Egyesült Államokban és Nagy-Britanniában végeztek az elmúlt évtizedekben. A számszerű eredmények szóródása viszonylag nagy (melynek okáról a későbbiekben még szót ejtünk), abban azonban a téma vezető közgazdászai általában megerősítenek bennünket, hogy a legjobb becslések szerint a statisztikai emberi élet értéke a 2-4 millió dolláros intervallumba esik.

zist igaznak fogadhatjuk el, akkor a módszer alkalmazása elől elhárulnak az akadályok, de az akkor is vizsgálat tárgya marad, hogy a számszerű eredmény valóban azt mutatja-e, amit várunk tőle.

Mindezek érdekében azt kell a korábbiakban bemutatott módon felépülő regressziós modellünk kapcsán statisztikai tesztnek kitenni, hogy a kockázatot magában foglaló magyarázó változóhoz tartozó koefficiens szignifikánsan különbözik-e nullától, s előjele pozitív-e.

Mivel a magyar jövedelmi és vagyoni viszonyok számottevő mértékben eltérnek az amerikai, illetve brit viszonyoktól (s feltételezhetően az emberek preferencia-térképeiben is jelentős különbségeket lehetne kimutatni), nem tartjuk érdemesnek előzetes találgatások megfogalmazását a magyarországi emberi élet értékének nagyságával kapcsolatban. Ilyen célú hipotézis felállítására tehát nem kerül sor. Azt mindazonáltal látni kell, hogy alaphipotézisünk igazként való elfogadása önmagában nem túl érdekes számunkra (bár a módszer alkalmazhatósága szempontjából kétségtelenül megnyugtató lenne), sokkal izgalmasabb a számszerű végeredmény nagysága, nevezetesen a magyarországi statisztikai emberi élet értéke.

A téma kapcsán egyéb hipotézisek is megfogalmazhatók, ezek tesztelésére azonban a hedonikus ármódszer alkalmazása nem nyújt lehetőséget, az contingent valuation megközelítéssel érhető el, így e helyen ezekre nem térünk ki.

A FELHASZNÁLT ADATOK ÉS A MODELLBE BEVONT VÁLTOZÓK

A regressziószámításhoz felhasznált adatok túlnyomó többsége a TÁRKI Háztartás Monitor 1998-as felvételéből származik. A megfelelő leválogatások után egy 1287 elemű, az 1996. évi mikrocenzusnak megfelelően súlyozott mintát kaptunk, amely reprezentálja az 1998 márciusában dolgozó, s főállású jövedelemmel rendelkező felnőtt lakosságot.

A modellbe bevont szinte valamennyi változó forrása a Monitor vizsgálat. Az itt megjelenő változókat azok kisebb-nagyobb módosításával nyertük. Egyetlen kivételként az elemzésünkben kulcsszerepet játszó munkahelyi kockázati tényező változóját állítottuk elő más forrásból, mégpedig a Központi Statisztikai Hivatal (KSH), az Országos Munkabiztonsági és Munkaügyi Főfelügyelőség (OMMF) és a Magyar Bányászati Hivatal (MBH) által gyűjtött és felvett adatok segítségével.¹²

Az elemzés változói

Jelen vizsgálatunkhoz hasonló számításokat, mint arra korábban már utaltunk, számos angolszász kutató végzett az elmúlt évtizedekben, több tucat féle változót alkalmazva modelljeikben.¹³ A változók egy részét a megszokott demográfiai paraméterek alkotják, másik részük speciálisan a végzett munkához, munkahelyhez, gazdasági ágazathoz kapcsolódik. Majdnem minden vizsgálatban szerepel a szakszervezeti tagság mint magyarázó

¹² A kockázati mutató megalkotásához létszám és balesetszám jellegű adatokra van szükség (részletesen lásd később). Előbbiek forrása a KSH. (Az 1994–96-os évekre: Foglalkoztatottság és kereseti arányok 1996–1997 (Munkaügyi adattár), KSH, Budapest, 1998, 20. old.; az 1997–98-as évekre: Főbb munkaügyi folyamatok 1998. I–IV. negyedév, Időszaki tájékoztató, KSH, 1999. március, 18. old.) Utóbbiak forrása az OMMF és az MBH kiadvány formában nem hozzáférhető statisztikai, kimutatásai. Ezúton szeretnék köszönetet mondani Tokaji Károlynának és Harcsa Istvánnak (KSH), Galló Sándornak (OMMF), valamint Lukács Györgynek (MBH), akik az adatgyűjtésben nagy segítséget nyújtottak.

¹³ Leggyakrabban alkalmazott magyarázó változók az alábbiak: iskolázottság, szakszervezeti tagság, rassz, foglalkozási kategória, kor, nem, munkában eltöltött idő, családi állapot s a régió.

tényező. Ezt Magyarországon értelmetlen jelenleg vizsgálni, sőt *Sandy-Elliott* (1996) szerint másutt is az, mivel a szakszervezetek általában sokkal inkább a kockázat csökkentése érdekében lépnek fel, semmint az érte járó ex ante kompenzáció kivívásáért.

A felsoroltakon kívül mintegy félszáz egyéb változó jelenik meg a tanulmányokban, némelyek (például *A. Marin* és *G. Psacharopoulos* (1982) tanulmányában a foglalkozás kívánatossági foka) a jövedelmi szint magyarázatában kifejezetten szokatlanok, s jellemző egyes változók szorzatának (például iskolázottság×kor, nem×családi állapot), illetve (például az iskolázottság, a céglétszám, a kor) négyzetének mint új változónak felvétele a modellbe.

A hagyományokon túl természetesen a Monitor vizsgálat által nyert adatok köre jelentette a korlátot a lehetséges magyarázó változók kiválasztásánál. A legtágabb modell az alábbi független változókat tartalmazza.¹⁴

HALKOCK – a halálos munkabaleset kockázata az adott ágazatban, (1/millió)/hónap mértékegységgel (részletesebben lásd később).

BALKOCK – a munkabaleset kockázata az adott ágazatban, (1/millió)/hónap mértékegységgel (részletesebben lásd később).

VAROS – város-e a lakóhely, kétértékű (1=igen, 0=nem).

BUDAPEST – Budapest-e a lakóhely, kétértékű (1=igen, 0=nem).

DUNANTUL – Dunántúl-e a lakóhely, kétértékű (1=igen, 0=nem).

HAZAS – házas-e, kétértékű (1=igen, 0=nem).

NEM – mi a neme, kétértékű (1=férfi, 0=nő).

KOR – kor években kifejezve.

ISKOLA – a teljesített legmagasabb iskolai végzettség szerinti évek számával kifejezett iskolázottság.

GYERMEK – a 18 év alatti háztartástagok száma.

IDNYELV – beszél-e valamilyen idegen nyelvet, kétértékű (1=igen, 0=nem).

BETEGSEG – van-e olyan betegsége, amire rendszeresen gyógyszer kell szednie, kétértékű (1=van, 0=nincs).

MELLEK – van-e mellékállása, kétértékű (1=van, 0=nincs).

MEZOGAZD – munkahelye a mezőgazdaság ágazatához tartozik-e, kétértékű (1=igen, 0=nem).

IPAR – munkahelye az ipar ágazatához tartozik-e, kétértékű (1=igen, 0=nem).

ONALLO – önálló, illetve vállalkozó-e, kétértékű (1=igen, 0=nem).

VEZSZELL – vezető és/vagy szellemi foglalkozású-e, kétértékű (1=igen, 0=nem).

CEGMERET – cégénél hányan dolgoznak, kétértékű (1=10-en vagy többen, 0=kevesebben).

ALLAMTUL – cége állami, illetve önkormányzati tulajdonban van-e, kétértékű (1=igen, részben vagy egészben, 0=nem).

MAGYTUL – cége teljes mértékben magyar tulajdonban van-e, kétértékű (1=igen, 0=nem, részben vagy egészben külföldi).

KIEGJUT – kapott-e élet-, nyugdíjbiztosítást, cégautó-, autófenntartási költségtérítést, benzinpénzt, vásárláshoz vagy szállításhoz kedvezményt, munkahelyi orvosi ellátást és kedvezményes üdülési lehetőségek közül legalább egyet az elmúlt 12 hónapban a munkahelyétől, kétértékű (1=igen, 0=nem).

MAGYAR – melyik népcsoporthoz tartozónak vallja magát, kétértékű (1=magyar, 0=egyéb).

CIGANY – melyik népcsoporthoz tartozónak vallja magát, kétértékű (1=cigány, 0=egyéb).

UJSAGOLV – rendszeresen olvas-e újságot, kétértékű (1=igen, 0=nem).

KONYVOLV – rendszeresen olvas-e könyvet, kétértékű (1=igen, 0=nem).

SZINHAZ – volt-e az elmúlt 12 hónapban színházban, kétértékű (1=igen, 0=nem).

HANGVERS – volt-e az elmúlt 12 hónapban hangversenyen, kétértékű (1=igen, 0=nem).

MUZEUM – volt-e az elmúlt 12 hónapban múzeumban, kétértékű (1=igen, 0=nem).

¹⁴ A Függelék F.2. táblája bemutatja a változók leíró statisztikáját. A változók operacionalizálása az eredeti vizsgálatban szereplő kérdésekre megfogalmazott válaszlehetőségek alapján történt. Mivel ezek többnyire zárt kérdések voltak, így az operacionalizálás is minden esetben egyértelmű volt.

SPORT – szokott-e sportolni, kétértékű (1=igen, 0=nem).

BARATOK – jó barátok száma.

SZELSPOL – politikailag szélsőségesnek véli-e magát, kétértékű (1=igen, 0=nem).

Rövidebb magyarázatot talán csak az utolsó nyolc változó felvétele igényel. Egyebek között *C. Brown* (1980) hozza fel kritikaként a területen végrehajtott kutatásokkal szemben, hogy azok, megfelelő adatok hiányában, figyelmen kívül hagynak nagyszámú, potenciálisan fontos személyes paramétert. Olyan tényezők tehát, amelyek az adott egyénre jellemzők, nem változnak az idővel, de nem tömöríthetők be a hagyományos változóba. *Brown* ennek kiküszöbölésére azt a megoldást választja, hogy nem keresztmetszeti, hanem longitudinális vizsgálatokat végez. *S. A. Low* és *L. R. McPheters* (1983), szintén e hatás kivédésére, más utat választ. Kutatásukat csupán a városi rendőrök körében végezték el, ezzel csökkentve az emberek kockázatészlelésének, képzettségének s egyéb számtalan jellemzőjének különbségéből adódó torzításokat, mivel e társadalmi csoport sok tekintetben viszonylag homogénnek tekinthető.

A változólistán szereplő utolsó nyolc változó felvételével nekünk is az volt a célunk, hogy az emberek különbözőségéből adódó torzításokat csökkentjük, azt valamilyen formában kontrolláljuk a regressziós egyenletben is. Olyan feltételezések húzódnak meg ennek hátterében, mint hogy például az ember olvasási szokása jó mutatója olyasféle belső tulajdonságnak, amely a jövedelem magyarázatát segíti, de nincs lefedve a többi változóval. Az ok-okozati viszonyok természetesen nem egyértelműek, mégis, jobbnak látjuk első körben „kipróbálni” ezeket a változókat.

A kockázati változók

A dolgozónak az adott munka (mely szempontunkból a munkahellyel és az ott végzett munkával jellemezhető) elfogadásával járó kockázata két dimenzióban jelentkezik.¹⁵ Egyrészt az életét, másrészt az egészségét is veszély fenyegeti. A statisztikai emberi életmérése szempontjából csupán az első típusú kockázat releváns, azonban a második típusú kockázat is befolyásolhatja a jövedelem nagyságát, ezért célszerű azt is szerepeltetni magyarázó változóként.¹⁶

Nem volt problémamentes a megfelelő HALKOCK és BALKOCK értékeknek a minta egyes elemeihez való hozzárendelése. Bár eredeti adatbázisunkban volt olyan változó, amely a válaszadó munkahelyének ágazati besorolását tartalmazta, az itt alkalmazott ka-

¹⁵ Hagyjuk most figyelmen kívül az olyan jellegű kockázatokat, mint például az állás elvesztésének veszélye, bár kétségtelen, hogy az ilyen adatok esetleges ismerete sokat javítana regressziós modellünkön.

¹⁶ Az életet fenyegető kockázat mértékét a HALKOCK változóval jellemeztük, amelynek értékeit a következőképpen állítottuk elő. Az OMMF minden esztendőben összesíti a halálos munkabalesetek számát a KSH elnökének érvényes közleménye alapján felállított ágazati bontásban. (Egyetlen kivétel a bányászat, mely ágazat baleseti statisztikái az MBH-tól származnak.) Ezen érték, s az alkalmazásban állóknak az adott ágazatban mért átlagos évi állományi létszámának (amelyet a KSH tesz közzé) hányadosával definiálható egyfajta ex post halálos kockázati mérték az adott évre. Ez természetesen évről évre változik, éppen ezért célszerűnek látszik valamiféle átlagos értékkel számolnunk. Mivel a Monitor vizsgálat 1998-ban készült, mi a megelőző öt év átlagával, vagyis az 1992–97-es adatok felhasználásával nyert mértékekkel dolgoztunk. Havi kockázatot számoltunk millió főre, mivel a kockázatnak a havi jövedelemre való hatására vagyunk kíváncsiak. Egy adott ágazatkategóriára tehát a HALKOCK változó azt mutatja, hogy az 1992 és 1997 között eltelt öt évben egymillió főből átlagosan hányan veszítették életüket munkabaleset következtében havonta. Ezt a kockázati mértéket tekintjük a pénzügyi kárpótlás alapjának az elvégzett regressziós számításnál. Nagyon hasonló módon definiáltuk a munkabalesetben elszenvedett sérülés kockázatának jellemzésére megalkotott BALKOCK változót is. A különbség mindössze annyi, hogy a hányados számlálójában nem a halálos balesetek száma, hanem az összes munkabalesetek száma szerepel. Itt kell megjegyezni, hogy a kockázat mérése a módszer kétségkívül egyik legkritikusabb pontja. A problémák forrásával foglalkozó részben erre részletesen kitértünk.

tegóriák némelyike nem bizonyult kellőképpen szűknek. Ezekben az esetekben a további alkategóriákba való besorolást a foglalkozás alapján végeztük.

Az eredményváltozó

Vizsgálatunk célja a marginális pénz–kockázat átváltás meghatározása, ezért függő változónknak valamiféle jövedelemadatnak kell lennie, méghozzá olyannak, amely logikailag ok-okozati összefüggésben van a munkahelyi kockázat nagyságával. Éppen ezért nem tartottuk célszerűnek a Monitor vizsgálat által feltárt havi nettó összjövedelem alkalmazását. Célunknak megfelelőbbnek látszott a főállásból származó jövedelemmel dolgozni. Az azonban kérdéses lehet, hogy érdemes-e ezen értékhez hozzáadni olyan jövedelmeket, mint a prémium, jutalom, a 13. havi fizetés, a nyereségrészesedés, a vállalkozásból származó nyereség, az osztalék, illetve részesedés. Ezek az elemek egyrészt bizonytalanok, vagyis nem tekinthetők biztos jövedelemnek, másrészt azonban a legtöbb esetben azért jól kalkulálhatók, esetleg előzetes megállapodás szerint kvázi-biztosan részét is képezik a munkáért járó kompenzációnak.

Ennek megfelelően kétféleképpen is elvégeztük a számításokat, a szűkebb jövedelemadatokat a JOVED1, a tágabbakat a JOVED2 változóba foglalva. A módszertani hagyományokat követve modellünket a jövedelmi változó természetes alapú logaritmus (LNJOVED1, illetve LNJOVED2 változók) mint eredményváltozó alkalmazásával is lefuttattuk.¹⁷

Felmerülhet a kérdés, hogy nettó vagy bruttó jövedelemadatokkal érdemes-e dolgozni. Ilyen szempontból szerencsénk volt, az irodalom ugyanis a számunkra is rendelkezésre álló nettó adatok használatát javasolja (*Moore–Viscusi*; 1988b. 379. old.).

A REGRESSZIÓSZÁMÍTÁS EREDMÉNYEI

Első körben tehát négy formában futtattuk le regressziós modellünket. A legfontosabb eredményeket az 1. tábla tartalmazza, melyek alapján az alábbi megállapításokat tehetjük.

a) Az *F*-próbafüggvényre nyert értékek alapján kijelenthető, hogy „globálisan nézve” mind a négy regressziófüggvény 0-tól szignifikánsan különböző hányadát magyarázza meg az eredményváltozónak ($F_{0,95}(31,1253) \approx 1,46$).

b) A második típusú jövedelmi változóra épített modell esetén nagyobb magyarázó erőt tapasztalunk, mint az elsőnél (nagyobb a szabadságfokkal korrigált determinációs együttható), mindkét típusú jövedelmi változó esetén azonban a logaritmikus forma beszédesebb modelltől árulkodik.

c) A viszonylag nagyszámú változóból csak kevés mutatkozik szignifikánsnak, ezért mindenképpen szükségesnek látszik modelljeinket tovább specifikálni.

d) A szokásos szinteken sem a HALKOCK, sem a BALKOCK változóhoz tartozó együttható nem mutatkozik szignifikánsnak.

e) Mivel a kockázat inferior jószág, vagyis a jövedelem növekedésével csökken az iránta való kereslet (*Jones–Lee*; 1976, *Thaler–Rosen*; 1976), így várható lenne, hogy a Pearson-féle korrelációs együttható értéke negatív lesz a jövedelem és a kockázat között. Más a helyzet a parciális korrelációs együtthatóval, amelynél már kiszűrjük egyéb tényezők hatását. Modelljeinkben szinte valamennyi esetben pozitív, de meglehetősen alacsony értékű korreláció mutatható ki. Ez számít talán a legmeglepőbb eredménynek.

¹⁷ *Marin–Psacharopoulos* (1982. 833. old.) szerint elméletileg indokoltabb a logaritmikus alak használata.

f) A statisztikai emberi élet értéke a 78-346 millió forintos tartományban helyezkedik el az egyes modelleknél.

1. tábla

Az első kör modelljei

Megnevezés	1.1. modell	1.2. modell	1.3. modell	1.4. modell
Függő változó	JOVED1	JOVED2	LNJOVED1	LNJOVED2
Magyarázó változók száma	31	31	31	31
5 százalékon szignifikáns magyarázó változók száma	8	8	12	14
Korrigált R^2	0,196	0,229	0,345	0,374
Globális F -érték	11,107	13,268	22,793	25,723
BALCKOCK koefficiens	0,426	-0,339	1,871E-05	1,184E-05
BALCKOCK koefficiens standard hibája	0,640	0,703	0,000	0,000
BALCKOCK koefficiens szignifikanciaszintje	0,506	0,630	0,059	0,051
BALCKOCK Pearson-korrelációja a függő változóval	0,030	0,025	0,058	0,050
BALCKOCK parciális korrelációja a függő változóval	0,019	-0,014	0,053	0,033
HALCKOCK koefficiens	138,033	346,492	2,077E-03	3,723E-03
HALCKOCK koefficiens standard hibája	121,934	133,963	0,002	0,002
HALCKOCK koefficiens szignifikanciaszintje	0,258	0,010	0,271	0,238
HALCKOCK Pearson-korrelációja a függő változóval	0,069	0,092	0,078	0,084
HALCKOCK parciális korrelációja a függő változóval	0,032	0,073	0,031	0,055
<i>Statisztikai emberi élet értéke (millió forint)</i>	<i>138,033</i>	<i>346,492</i>	<i>77,99</i>	<i>146,124</i>

A második körben backward eljárást alkalmazva igyekeztünk kiszűrni a nem szignifikáns változókat (5 százalékos küszöböt alkalmazva) és lépni az optimális modell irányába. A kapott eredményeket a 2. tábla tartalmazza.

2. tábla

A második kör modelljei

Megnevezés	2.1. modell	2.2. modell	2.3. modell	2.4. modell
Függő változó	JOVED1	JOVED2	LNJOVED1	LNJOVED2
Magyarázó változók száma	11	12	13	14
5 százalékon szignifikáns magyarázó változók száma	11	12	13	14
Korrigált R^2	0,197	0,227	0,343	0,372
Globális F -érték	29,577	32,437	52,650	55,402
BALCKOCK koefficiens	–	–	2,938E-05	–
BALCKOCK koefficiens standard hibája	–	–	0,000	–
BALCKOCK koefficiens szignifikanciaszintje	–	–	0,000	–
HALCKOCK koefficiens	236,075	369,320	–	5,978E-03
HALCKOCK koefficiens standard hibája	89,827	98,898	–	0,001
HALCKOCK koefficiens szignifikanciaszintje	0,009	0,000	–	0,000
<i>Statisztikai emberi élet értéke (millió forint)</i>	<i>236,075</i>	<i>369,320</i>	<i>–</i>	<i>234,896</i>

Kiszámítottuk továbbá a HALCKOCK és BALCKOCK változók közötti Pearson-féle korrelációs együtthatót, amely 0,664-nek adódott, vagyis 44,1 százalékban magyarázza egymás varianciáját a két változó. (Egy-egy regressziószámítás futtatásával megállapított

tuk, hogy a többi magyarázó változóval együttesen 53,4 százalékban magyarázza a BALKOCK változó a HALKOCK változót, míg fordítva ez az érték 47,4 százalék.)

Mindezek figyelembevételével az alábbi megállapítások látszanak helyénvalónak.

a) A HALKOCK és a BALKOCK változó nagymértékű korrelációja jelentős multikollinearitást visz a modellbe, ezért célszerű csak a HALKOCK változót meghagyni. Mindezt alátámasztják az eredmények is, mind a négy modellben egyszerre csupán az egyik változó mutatkozott szignifikánsnak, az viszont, miután a másik változó kikerült a modellből, feltűnően alacsonyabb szinten, mint eredeti modelljeinkben. A jelenség hátterében az állhat, hogy a két változó valójában majdnem ugyanazt méri, mivel az egyes ágazatokban a relatív halálos és nem halálos kockázatok igen nagy hasonlóságot mutatnak. Szándékainknak sokkal inkább megfelel a HALKOCK változó modellben való szerepeltetése, tudnunk kell azonban, hogy a BALKOCK változó elhagyása után, éppen a hasonlóság miatt, a HALKOCK változóhoz tartozó koefficiensből nyerhető kompenzációs összegben már a nem halálos kockázatokért járó kompenzáció is szerepel, még ha annak mértéke jóval (hózzávetőleg két nagyságrenddel) csekélyebb is a halálos kockázatért járó kompenzáció mértékénél.

b) A korrigált többszörös determinációs együtthatók értékei alig csökkentek a kiterjesztett modellekhez képest, amiből az a következtetés vonható le, hogy modelljeink magyarázó ereje gyakorlatilag nem romlott a nagyszámú változó kiesésével.

c) A statisztikai emberi élet értéke a 235-369 millió forintos intervallumba adódik.

d) Most is a negyedik típusú, vagyis az LNJOVED2 magyarázó változóval dolgozó modell mutatkozik optimálisnak, ezért a továbbiakban már csak ezzel foglalkozunk.

A 3. tábla a 2.4. modell magyarázó változóihoz tartozó együtthatókat, azok standard hibáját és szignifikanciaszintjét mutatja be. Valójában csupán a HALKOCK változóhoz tartozó koefficiens tart számot érdeklődésünkre, a többi változó leginkább a legjobban illeszkedő modell kialakítása szempontjából érdekes számunkra.

3. tábla

A 2.4. modell magyarázó változói

Változó	Együttható	Standard hiba	Szignifikancia-szint
(Konstans)	9,058	0,082	0,000
ISKOLA	4,825E-02	0,006	0,000
NEM	0,192	0,025	0,000
VEZSCELL	0,182	0,032	0,000
BUDAPEST	0,221	0,033	0,000
KIEGJUT	0,132	0,024	0,000
KOR	6,180E-03	0,001	0,000
CEGMERET	0,199	0,034	0,000
ONALLO	0,264	0,053	0,000
UJSAGOLV	0,108	0,025	0,000
HALKOCK	5,978E-03	0,001	0,000
MAGYTUL	-8,930E-02	0,025	0,000
GYERMEK	3,102E-02	0,012	0,011
VAROS	6,132E-02	0,026	0,019
MUZEUM	5,744E-02	0,028	0,042

Ahogy az várható volt, az iskolai végzettség, a nem, a vezető, illetve szellemi foglalkozás, a budapesti és a városi lakóhely és az életkor szignifikánsan, pozitív előjellel befolyásolja a jövedelem nagyságát. Nem volt talán ennyire egyértelmű a cégmértet, az

önállóság, s a gyermekek számának pozitív, illetve a magyar tulajdonú cég negatív irányú hatása. Végül fontos látnunk, hogy a kiegészítő juttatások létének ténye a jelzett pozitív irányú kapcsolat esetén nem állhat ok-okozati összefüggésben a jövedelem nagyságával, hiszen nem mondható, hogy az emberek fizetése szisztematikusan nő az esetleges kiegészítő juttatások léte miatt. (Negatív kapcsolat esetén megfogalmazhatnánk ok-okozati összefüggést is.) Ezt a változót tehát ki kell vennünk a modelltől. Az ily módon felírt 3.1. modell néhány fontos jellemzőjét tartalmazza a 4. tábla.

Két egyéb felírással próbálkoztunk. Elsőként a 3.1. modell magyarázó változói közé felvettük az ISKOLA×ISKOLA, illetve KOR×KOR változókat (3.2. modell, 4. tábla). Az ilyen jellegű interakcióknak az egyes vizsgálatokban való használatára a korábbiakban már történt utalás. A forward eljárással lefutott regressziós modellben mindkét változó szignifikánsnak mutatkozott, ezúttal azonban „elvesztette szignifikanciáját” az ISKOLA és a GYERMEK változó. Végezetül a HALKOCK változón kívül a 3.1. modell valamennyi változójának mindegyik változóval való keresztszorzatát mint új változót felvettük a szintén forward eljárással lefutott modellt (3.3. modell, 4. tábla). Ez $((12 \times 12) / 2 - 3 =)$ 69 új változót jelentett, hiszen a kétértékű változók önmagukkal való szorzata az eredeti változót eredményezi. Fontos felhívni a figyelmet arra, hogy ez a felírás sokkal inkább tekinthető egy kis játéknak, mintsem módszertanilag alaposan alátámasztott kezdeményezésnek. A 82 változóból 24 bizonyult szignifikánsnak, s bár a többszörösen korrigált determinációs együttműködő e modell esetén jelzi a legnagyobb értéket, a modell „tisztatlansága” miatt csak nagy óvatossággal szabad kezelni az így nyert eredményeket.

4. tábla

A harmadik kör modelljei

Megnevezés	3.1. modell	3.2. modell	3.3. modell
Függő változó	LNJOVED2	LNJOVED2	LNJOVED2
Magyarázó változók száma	13	13	24
5százalékon szignifikáns magyarázó változók száma	13	13	24
Korrigált R^2	0,358	0,368	0,404
Globális F -érték	56,087	58,477	37,326
HALKOCK koefficiens	6,145E-03	6,399E-03	9,980E-03
HALKOCK koefficiens standard hibája	0,001	0,001	0,005
HALKOCK koefficiens szignifikanciaszintje	0,000	0,000	0,034
<i>Statisztikai emberi élet értéke (millió forint)</i>	<i>241,478</i>	<i>251,492</i>	<i>392,935</i>

Számításaink alapján a következőket állapíthatjuk meg. A statisztikai emberi élet értéke 1998-ban Magyarországon nagy valószínűséggel a 78-393 millió forintos intervallumban helyezkedik el, legmegbízhatóbbnak a 250 millió forint körüli értékek tekinthetők.

A MÓDSZER KORLÁTJAI – LEHETSÉGES HIBAFORRÁSOK

A regressziószámítás alkalmazásánál megszokott, az eljárás matematikai statisztikai jellegzetességeiből adódó hibaforrásokra e helyütt nem térünk ki. Fontosabbnak érezzük annak a kérdésnek a vizsgálatát, hogy vajon a felállított elméleti modell megfelelően jól

írja-e, írhatja-e le a valóságot. E fejezetben az ezzel kapcsolatos legfontosabb problémákat vesszük sorra.

A listát érdemes a leggyakrabban elhangzó kritikával kezdeni. Eszerint a munkaerőpiaci viszonyokat manapság egyáltalán nem versenytényezők alakítják, így nem is várható el, hogy olyan tényező, mint a kockázat, megjelenjék a munkabérekben. Elég, ha megnézzük, hogy a legkockázatosabb foglalkozások (például magasban dolgozó munkások, bányászok stb.) egyáltalán nem jól fizetettek, a legnagyobb keresetek sokkal inkább a szinte kockázat nélküli állásokban (például bankokban, tanácsadó cégeknél) érhetők el. E kritikát még könnyű kivédeni, hiszen a regressziószámítás lényege éppen abban áll, hogy kiszűrjük a jövedelem nagyságát alakító egyéb tényezők hatását (ezek felelősek például a bankár és a bányász jövedelmében megmutatkozó különbségért), s az ezek után fennmaradó eltérések magyarázatára használjuk a kockázat mértékét (*Marin–Psacharopoulos; 1982*).

Itt kell mindenképpen kitérnünk *R. Thaler* és *S. Rosen* (1976) tanulmányára, mely az első jelentős ilyen irányú kutatás tapasztalatait összegezi, s mindmáig az első számú hivatkozási alap e témakörben. A szerzők egyik fontos megállapítása, hogy a magasabb kockázattal rendelkező foglalkozásokat általában kevésbé kockázatkerülő emberek üzik, így esetükben a fizetési hajlandósági érték kisebb, ami alacsonyabb emberiélet-értéket eredményez, mint kevésbé kockázatos területeken. Ezt az eredményt *C. Olson* (1981) empirikusan is alátámasztotta. Vagyis az alkalmazott mintának reprezentatívnak kell lennie a foglalkozások, munkahelyek kockázatára nézve, másként nem nyerhető torzítatlan eredmény. (Thaler és Rosen kifejezetten magas kockázatú foglalkozásokat vizsgált.)

Náluk került elő elsőként az a probléma is, hogy milyen jellegű kockázati adatokkal dolgozzunk. Két alapvető lehetőség mutatkozik: foglalkozási vagy iparági adatokkal. Az általuk használt foglalkozáshoz kötődő adatokat ráadásul úgy állították elő, hogy az adott foglalkozást űzők körében kiszámolták a statisztikailag várható halálesetek számát, majd ezt kivonták a tényleges értékekből, ezzel megkapva az extra halálozási számokat, amit tehát az adott foglalkozásnak tulajdoníthatunk. E módszert sok bírálat érte (például *Viscusi; 1979, Lipsey; 1976*), mondván, az így előállított extra kockázat nemcsak az adott foglalkozás veszélyességével, hanem a foglalkozást űzők egyes jellegzetes tulajdonságai, életstílusjegyeivel, jövedelmi viszonyaival stb. is szoros kapcsolatban van.¹⁸

Elméletileg kizárólag a dolgozók által észlelt kockázat mértékének vizsgálata lenne helyes, ilyen adatok hiányában azonban a kutatások túlnyomó többsége valós kockázati adatokra támaszkodik, azzal a nem éppen realiztikus feltételezéssel élve, hogy ez jól helyettesíti a megfelelő észlelt mértékeket.¹⁹ Kimutatott azonban az a tendencia, hogy az emberek a kismértékű kockázatokat (s a munkahelyi kockázatok az esetek többségében feltétlenül ide sorolandók) túl-, a nagymértékűeket pedig alulbecsülik (*Moore–Viscusi; 1988a. 482. old.*). Önmagában ez a hatás felfelé torzítja a statisztikákból számolt emberi élet értékét, nem tudhatjuk azonban, hogy a kockázateszlelésben megmutatkozó egyéb

¹⁸ Így lehetséges azután, hogy például a pincérek foglalkozását az adatok háromszor olyan veszélyesnek tüntetik fel, mint a tűzoltókét (*Thaler–Rosen; 1976. 288. old.*).

¹⁹ *Gegax–Gerking–Schulze* (1991) észlelt kockázatokkal végezte el vizsgálatait. Az így nyert eredmények gyökeresen eltértek a kockázati statisztikákon alapuló számítások végeredményeitől. Úgy találták, hogy a konvencionális kockázati mértékek ugyan szignifikánsan korrelálnak az észlelt értékekkel, ez utóbbi varianciájának azonban csupán 9 százalékát magyarázzák meg. Ezzel kissé ellentétes megállapításra jutott *Viscusi* (1979. 246. old.), aki vizsgálatában nem talált szignifikáns különbséget az észlelt és tényleges kockázati adatokon alapuló számítások eredményei között.

tökéletlenségek, azok különbségei az egyes foglalkozások, társadalmi csoportok stb. között hogyan befolyásolják elemzésünket.

Dillingham (1985) teljes kutatását annak szentelte, hogy kimutassa, vajon a kockázati változó megválasztása milyen hatást gyakorol az emberi élet értékének meghatározására. A szerző azt tapasztalta, hogy a korábbi kutatások eredményeinek jelentősnek tekinthető szóródását nagy részben magyarázza a használt kockázati definíció. Egyrészt az iparági adatokon alapuló becslések szignifikánsan nagyobbak a foglalkozási adatokon alapulóknál, másrészt a biztosítási matematika eszközeivel meghatározott pótlólagos halálzási arányon alapuló számítások szignifikánsan alacsonyabb értéket eredményeznek a munkabaleseti statisztikákon alapulóknál. Mivel mi számításainkat iparági bontásban rendelkezésre álló munkabaleseti statisztikák segítségével végeztük, így várható, hogy e tény a kapott eredmény felfelé való torzítottságában jelentkezik. Empirikus eredményeivel *Leigh* (1987) is alátámasztotta a kockázati változó kiválasztásának rendkívüli jelentőségét.

Sajnálatos módon azonban az általunk használt munkabaleseti adatok még statisztikai szempontból is tökéletlenek, alulbecsülik a kockázati mértékeket, ami az emberi élet értéket pozitív irányba torzítja. Ennek egyik oka az, hogy eleve csak a munkaidőben keletkező balesetekkel foglalkozik, márpedig feltételezhetően igen sok munkához kötődő baleset történik munkaidőn kívül, illetve sok baleset, megbetegedés és haláleset gyökere nyúlhat vissza a korábbi munkahelyhez, foglalkozáshoz (például a vegyiparban), amelyek nyilvánvalóan szintén nem jelennek meg az adott területről készült statisztikákban. A másik ok az – s ebben az OMMF munkatársai egyértelműen megerősítettek –, hogy a cégek különböző okokból kifolyólag, erősen érdekelték abban, hogy a bekövetkezett, sokszor halálos munkabaleseteket ne vegyék jegyzőkönyvbe, hanem privát módon egyezzenek meg a sérülttel, vagy annak családjával. Sok esetben csak ennek kudarca esetén fordulnak a hivatalos eljárási módhoz.

Herzog-Schlottmann (1990) szerint a hedonikus ármódszert azért kell nagy óvatossággal kezelnünk, mert arra a valóságban nem teljesülő alapfeltételezésre épít, hogy a kockázatsökkenésért való fizetési határhajlandóság egyenlő annak határkölségével. A szerzőpáros szerint a valóságban e fizetési hajlandósági értékek nagyobbak, mint a fizetésben megjelenő kompenzációs érték (mely ily módon sokkal inkább a határkölséget tükrözi), aminek oka a tökéletlen információ, a nem hatékony alkufolyamat és a tranzakciós költségek. Ezek a jelenségek Magyarországon feltehetően még hatványozottabban érvényesülnek. Mindez lefelé torzítja a kapott emberiélet-értéket.

Végül *Moore* és *Viscusi* (1988b), illetve *Arnould* és *Nichols* (1983) azt rója fel a korábbi kutatásoknak, hogy azok nem veszik figyelembe az esetleges balesetek esetén járó, általában valamilyen biztosításból származó ex post kompenzációt. Ez szintén negatív irányba hat torzítólag az emberi élet értékére nézve.

*

Kutatásunk alaphipotézise az volt, hogy a munkavállalók által végzett munka kockázatának mértéke pozitív irányban befolyásolja az érte kapott munkabér nagyságát. E hipotézisünket az empirikus teszt alapján igaznak fogadhatjuk el, mivel az „optimális” regressziós modellben a jövedelem nagyságát magyarázó kockázati változóhoz tartozó koeficiens pozitívnak, s nullától minden konvencionális szignifikanciaszinten különbözőnek

bizonyult. Nem alaptalan tehát a hedonikus ármódszerrel próbálkoznunk a sajátosan definiált statisztikai emberiélet-érték meghatározásakor Magyarországon.

A valamilyen eredményt produkáló 10 modell a 78-393 millió forintos intervallumban jelöli ki ezt az értéket, legmegbízhatóbbnak a 250 millió 1998-as forint körüli összeg tekinthető. Eredményeink helyes értelmezése, s a megfelelő konzekvenciák levonása – figyelembe véve az előző fejezetben bemutatott torzító hatásokat – egyszerre követel óvatosságot és bátorságot. Ne feledjük, a statisztikai emberi élet értékének megállapítására a közszektorban történő beruházások értékelése, értékelhetősége miatt volt szükségünk. Továbbra sem tudjuk, hogy ha egy program egy statisztikai emberi életet 110 vagy 350 millió forintos költségen képes megmenteni, akkor annak „futása” találkozik-e a társadalom helyeslésével. Ilyen mértékű pontosságra nem adnak eredményeink okot. Az azonban kimondható, hogy ha például 25 millió forint ez az érték, akkor biztosan „igen”, míg 2,5 milliárd forint esetén biztosan „nem” a társadalom válasza.

Eredményeinkbe vetett bizalmunkat növelheti, hogy azok harmonizálnak az angolszász kutatások során nyert eredményekkel. A nagyjából egymillió dollárnak megfelelő érték kissé alulmarad a legmegbízhatóbb amerikai és angol értékekhez képest, e tény azonban semmiképpen sem szabad úgy magyaráznunk, tekintenünk, hogy a magyar statisztikai emberi élet kevesebbet ér, mint például az amerikai, jóllehet a számok gyanútlan összevetése ezt sugallná. A statisztikai emberi élet ugyanis csupán egy adott populációban értelmezhető, így közvetlen összehasonlításra nem alkalmas. Törvényszerű, hogy a nominálisan kevesebb jövedelem és vagyon felett diszponáló közösségek kevesebbre értékeljék saját életüket is. Ez az érték azonban sokkal inkább az adott közösséget jellemzi, semmint a benne élő egyes egyéneket. Ha szemléletesen akarunk fogalmazni, azt mondhatjuk, hogy ha egy amerikai hazánkba látogat, akkor az ő élete ugyanazon értékelés alá esik, mint a közösség magyar tagjaié, és viszont.

Dolgozatom célja az empirikus vizsgálatok mellett az is volt, hogy tágabb betekintést engedjek abba a problémafelvetés- és gondolkodásmódba, ami korunk közgazdászait, legalábbis egy részüket jellemzi. Az emberi élet értékelésével a közgazdaságtan legfőbb céljának tesz eleget: a társadalmi problémák megoldásában kíván segédkezni.

FÜGGELÉK

F.1. A legfontosabb kutatások eredményeinek összefoglalása

Tanulmány	Az emberi élet értéke (ezer dollár, 1990, legjobb becslés)
Munkaerő-piaci megközelítés	
Thaler–Rosen (1976) – USA	1 000
Brown (1980) – USA	1 500
Marin–Psacharopoulos (1982) – U.K.	2 800
Smith (1983) – USA	1 300
Arnould–Nichols (1983) – USA	900
Low–McPheters (1983) – USA	1 200
Moore–Viscusi (1988a) – USA	7 300
Gegax–Gerking–Schulze (1991) – USA	1 600

(A tábla folytatása a következő oldalon.)

(Folytatás.)

Tanulmány	Az emberi élet értéke (ezer dollár, 1990, legjobb becslés)
Egyéb piaci megközelítés	
Ghosh–Lees–Seal (1975) – U.K.	700
Dardis (1980) – USA	600
Portney (1981) – USA	800
Ippolito–Ippolito (1984) – USA	700
Atkinson–Halvorsen (1990) – USA	4 000
Contingent valuation megközelítés	
Acton (1973) – USA	100
Melinek–Woolley–Baldwin (1973) – U.K.	170
Jones–Lee (1976) – U.K.	15 600
Jones–Lee–Hammerton–Philips (1985) – U.K.	3 700
Viscusi–Magat–Huber (1991) – USA	9 700

Forrás: Fisher–Chestnut–Violette (1989), Table 1, 90. old., Jones–Lee (1989), Table 2.2–2.3, 91–93. old., Viscusi (1993), Table 2, 5, 6, 1926–1927., 1936., 1940. old. (Az F.1. tábla az egyes kutatásoknak a fenti szerzők szerinti legjobb becslését tartalmazza a statisztikai emberélet értékéről, 1990. decemberi dollárértéken megadva. Ilyen formában azonban csupán Viscusi (1993) tanulmányában szerepelnek az adatok, a másik két forrásból vett adatokat igyekeztünk – a tapasztalt matematikai jellegű ellentmondások ellenére – a lehető legpontosabban konvertálni.)

F.2. Az egyes változók átlaga és szórása

Változó	Átlag	Szórás
Függő változók		
JOVED1	37 511,688	29 620,518
JOVED2	39 176,085	33 218,324
LNJOVED1	10,389	0,508
LNJOVED2	10,419	0,526
Független változók		
HALKOCK	6,875	8,899
BALKOCK	1 189,898	1 595,225
VAROS	0,659	0,483
BUDAPEST	0,180	0,391
DUNANTUL	0,352	0,487
HAZAS	0,644	0,488
NEM	0,514	0,509
KOR	38,288	11,007
ISKOLA	11,357	2,563
GYERMEK	0,829	0,966
IDNYELV	0,186	0,397
BETEGSEG	0,202	0,409
MELLEK	0,070	0,261
MEZOGAZD	0,067	0,254
IPAR	0,372	0,493
ONALLO	0,075	0,268
VEZSZELL	0,345	0,486
CEGMERET	0,805	0,404
ALLAMTUL	0,352	0,486
MAGYTUL	0,677	0,476
KIEGJUT	0,446	0,506

(A tábla folytatása a következő oldalon.)

(Folytatás.)

Változó	Átlag	Szórás
MAGYAR	0,974	0,164
CIGANY	0,011	0,108
UJSAGOLV	0,651	0,486
KONYVOLV	0,297	0,465
SZINHAZ	0,231	0,429
HANGVERS	0,112	0,322
MUZEUM	0,331	0,479
SPORT	0,298	0,469
BARATOK	2,235	3,770
SZELSPOL	0,085	0,284

IRODALOM

- ACTON, J. P. (1973): *Evaluating public programs to save lives: the case of heart attacks*. Research Report R-73-02. The Rand Corporation, Santa Monica.
- ACTON, J. P. (1976): Measuring the monetary value of lifesaving programs. *Law and Contemporary Problems*, 40 évf. 4. sz. Autumn, 46–72. old.
- ARNOULD, R. J. – NICHOLS, L. M. (1983): Wage-risk premiums and workers' compensation: a refinement of estimates of compensating wage differential. *Journal of Political Economy*, 91. évf. 2. sz. 332–340. old.
- ATKINSON, S. E. – HALVERSON, R. (1990): The valuation of risks to life: evidence from the market for automobiles. *The Review of Economics and Statistics*, 72. évf. 1. sz. February, 133–136. old.
- BARA Z. – SZABÓ K. (szerk.) (1997): *Összehasonlító gazdaságtan*. Aula Kiadó, Budapest.
- BOARDMAN, A. E. – GREENBERG, D. H. – VINING, A. R. – WEIMER, D. L. (1996): *Cost-benefit analysis: concepts and practice*. Prentice Hall, USA, New Jersey.
- BROOME, J. (1978): Trying to value a life. *Journal of Public Economics*, 9. sz. 91–100. old.
- BROWN, C. (1980): Equalizing differences in the labor market. *The Quarterly Journal of Economics*, February, 113–134. old.
- CSONTOS L. – KORNAI J. – TÓTH I. GY. (1996): Az állampolgár, az adók és a jóléti rendszer reformja. Egy kérdőíves felmérés tapasztalatai. *Századvég*, Új Folyam, 2. sz. 3–28. old.
- DARDIS, R. (1980): The value of a life: new evidence from the marketplace. *The American Economic Review*, 70. évf. 5. sz. December, 1077–1082. old.
- DILLINGHAM, A. (1985): The influence of risk variable definition on value-of-life estimates. *Economic Inquiry*, 23. évf. 2. sz. April, 277–294. old.
- DOWNES, A. (1990): Politikai cselekvés a demokráciában: egy racionális modell. *Közgazdasági Szemle*, 37. évf. 9. sz. 993–1011. old.
- DUBLIN, L. I. – LOTKA, A. J. (1930): *The money value of a man*. Ronald Press, New York.
- FISHER, A. – CHESTNUT, L. G. – VIOLETTE, D. M. (1989): The value of reducing risks of death: a note on new evidence. *Journal of Policy Analysis and Management*, 8. évf. 1. sz. 88–100. old.
- GEGAX, D. – GERKING, S. – SCHULZE, W. (1991): Perceived risk and the marginal value of safety. *The Review of Economics and Statistics*, 73. évf. 4. sz. November, 589–596. old.
- GHOSH, D. – LEES, D. – SEAL, W. (1975): Optimal motorway speed and some valuation of time and life. *The Manchester School of Economic and Social Studies*, 43. évf. June, 134–143. old.
- HERZOG, H. W., JR. – SCHLOTTMANN, A. M. (1990): Valuing risk in the workplace: market price, willingness to pay, and the optimal provision of safety. *The Review of Economics and Statistics*, 72. évf. 3. sz. August, 463–470. old.
- IPPOLITO, P. M. – IPPOLITO, R. A. (1984): Measuring the value of life saving from consumer reactions to new information. *Journal of Public Economics*, 25. sz. 53–81. old.
- JONES-LEE, M. W. (1976): *The value of life*. An Economic Analysis. Martin Robertson, London.
- JONES-LEE, M. W. (1989): *The economics of safety and physical risk*. Basic Blackwell, Oxford.
- JONES-LEE, M. W. – HAMMERTON, M. – PHILIPS, P. R. (1985): The value of safety: results of a national sample survey. *The Economic Journal*, 95. sz. 49–72. old.
- KELMAN, S. (1981): Cost-benefit analysis. An ethical critique. *Regulation*, January/February, 33–40. old.
- LEIGH, J. P. (1987): Gender, firm size, industry, and estimates of the value-of-life. *Journal of Health Economics*, 6. évf. 3. sz. September, 255–273. old.
- LIPSEY, R. E. (1976): Comments on „the value of saving a life: evidence from the labor market”. In: TERLECKYI, N. E. (szerk.), *Household production and consumption*. Cambridge: NBER, 301–302. old.
- LOW, S. A. – MCPHETERS, L. R. (1983): Wage differentials and risk of death: an empirical analysis. *Economic Inquiry*, 21. sz. 271–280. old.
- MARIN, A. – PSACHAROPOULOS, G. (1982): The reward for risk in the labor market: evidence from the United Kingdom and a reconciliation with other studies. *Journal of Political Economy*, 90. évf. 4. sz. 827–853. old.

- MELINEK, S. J. – WOOLLEY, S. K. D. – BALDWIN, R. (1973): *Analysis of a questionnaire on attitudes to risk: Fire Research Note 962*. Joint Fire Research Organisation, Borehamwood.
- MISHAN, E. J. (1971): Evaluation of life and limb: a theoretical approach. *Journal of Political Economy*, 79. évf. 4. sz. 687–705. old.
- MISHAN, E. J. (1981): The value of trying to value a life. *Journal of Public Economics*, 15. évf. 1. sz. 133–137. old.
- MOORE, M. J. – VISCUSI, W. K. (1988a): Doubling the estimated value of life: results using new occupational fatality data. *Journal of Policy Analysis and Management*, 7. évf. 3. sz. 476–490. old.
- MOORE, M. J. – VISCUSI, W. K. (1988b): The quantity-adjusted value of life. *Economic Inquiry*, 26. évf. July, 369–388. old.
- MOORE, M. J. – VISCUSI, W. K. (1990): Discounting environmental health risks: new evidence and policy implications. *Journal of Environmental Economics and Management*, 18. sz. S51–S62. old.
- OLSON, C. (1981): An analysis of wage-differentials received by workers on dangerous jobs. *Journal of Human Resources*, 16. évf. 2. sz. Spring, 167–185. old.
- PORTNEY, P. R. (1981): Housing prices, health effects, and valuing reductions in risk of death. *Journal of Environmental Economics and Management*, 8. sz. 72–82. old.
- SANDY, R. – ELLIOTT, R. F. (1996): Unions and risk: their impact on the level of compensation for fatal risk. *Economica*, 63. évf. May, 250. sz. 291–309. old.
- SCHELLING, T. C. (1968): The life you save may be your own. In: CHASE, S.B. (szerk.), *Problems in public expenditure analysis. Papers presented at a conference of experts held September 15-16, 1966*. Washington, Brookings, 127–176. old.
- SMITH, A. (1992): *A nemzetek gazdagsága. E gazdagság természetének és okainak vizsgálata*. Közgazdasági és Jogi Könyvtáradó, Budapest.
- SMITH, V. K. (1983): The role of site and job characteristics in hedonic wage models. *Journal of Urban Economics*, 13. évf. May, 3. sz. 296–321. old.
- TÁRKI Adatbank. 1999. „TÁRKI Háztartás Monitor 1998 adatbázis.” (1.0A. CD-ROM verzió) Társadalomkutatási Informatikai Egyesülés, Budapest.
- THALER, R. – ROSEN, S. (1976): The value of saving a life: evidence from the labor market. In: TERLECKYI, N.E. (szerk.), *Household Production and Consumption*. NBER, Cambridge, 265–298. old.
- ULPH, A. (1982): The role of ex ante and ex post decisions in the valuation of life. *Journal of Public Economics*, 18. sz. 265–276. old.
- VISCUSI, W. K. (1978): Labor market valuations of life and limb: empirical evidence and policy implications. *Public Policy*, 26. évf. 3. sz. Summer, 359–386. old.
- VISCUSI, W. K. (1979): *Employment hazards. An investigation of market performance*. Harvard University Press, Cambridge, Massachusetts, London, England.
- VISCUSI, W. K. (1986): The Valuation of risks to life and health: guidelines for policy analysis. In: BENTKOVER, J. D. – COVELLO, T. – MUMPOWER, J. (szerk.), *Benefits assessment*. The State of the Art. D. Reidel, Dordrecht, 193–210. old.
- VISCUSI, W. K. (1992): *Fatal tradeoffs. Public and private responsibilities for risk*. Oxford University Press, New York, Oxford.
- VISCUSI, W. K. (1993): The value of risks to life and health. *Journal of Economic Literature*, 31. évf. December, 1912–1946. old.
- VISCUSI, W. K. – MAGAT, W. A. – HUBER, J. (1991): Pricing environmental health risks: survey assessments of risk-risk and risk-dollar trade-offs for chronic bronchitis. *Journal of Environmental Economics and Management*, 21. sz. 32–51. old.
- VISCUSI, W. K. – MOORE, M. J. (1989): Rates of time preference and valuation of the duration of life. *Journal of Public Economics*, 38. sz. 297–317. old.

SUMMARY

Setting a numerical value to the human life does not belong to the classical tasks of economics, not even to the popular pastime. The lives of people are really not commodities, not goods, do not form a part of the consumer's bundle. However, in those countries where the usage of modern methods – like cost-benefit analysis – in public policy goes back to decades these processes are quite common.

Economists of course don't have the courage and intention to estimate the value of the life of really existing people. According to the consensus among the leading economists in this field we can only determine the value of a specially defined statistical human life based upon the principle of willingness-to-pay of people.

To determine the value of a Hungarian statistical life a regression analysis-based approach is used in the study, the hedonic price method. The numerical results spread in the HUF 78-393 million interval, the most reliable value is somewhere around HUF 250 million (based on the value of the Hungarian currency in 1998), which is approximately USD one million. This number is a bit smaller than the results of the most competent Anglo-Saxon researches for the value of a statistical life in their country (they converge in the USD 2-4 million interval), however, we shouldn't conclude that the value of a Hungarian life is worth less than an American one. These values scilicet lean strongly on various parameters of the given population (the wealth and income relations etc.), that parameters differ across nations.

AZ ÁLTALÁNOSÍTOTT LINEÁRIS MODELL ÉS BIZTOSÍTÁSI ALKALMAZÁSAI

ROGER GRAY – KOVÁCS ERZSÉBET

A biztosítási károk alakulásának modellezésére jól alkalmazható az általánosított lineáris modell, amely alkalmas arra, hogy a nem normális eloszlást követő változók, például halál-
zási adatok, vagy a gépjárműtörés-károk várható számát becsüljük. A magyarázó változók között szerepeltethetünk mennyiségi és minőségi változókat, sőt ezek kölcsönhatásai is be-
építhetők a modellbe. A függő változó és a magyarázó változók prediktora között sajátos
függvénykapcsolat (link) létezhet. Ezek közül ismertet néhányat a tanulmány, és példákat
mutat a logit modell illesztésére, valamint tesztelésére.

TÁRGYSZÓ: Általánosított lineáris modell. Logit modell illesztése.

Az általánosított lineáris modell (Generalized Linear Model – GLM) kevert mérési
skálájú változók lineáris modellezésére alkalmas. Egyik speciális esete¹ a lineáris reg-
resszió, amelyben a normális eloszlást követő Y_i függő változó várható értéke $E(Y_i) = \mu_i$.
Az egyetlen magyarázó változót tartalmazó modell a következő alakban írható fel:

$$Y_i = \mu_i + \varepsilon_i, \quad \text{ahol } \mu_i = a + bx_i \text{ és } Y_i \sim N(\mu_i, \sigma^2). \quad /1/$$

Az $a + bx$ tagot az Y változó *lineáris prediktorának* nevezzük. Ez a modell alapvetően
két okból lehet nem kielégítő: ha az Y függő változó nem követ normális eloszlást, és ha
a függő változó függvénye a lineáris prediktornak, és nem közvetlenül egyenlő vele.

A normális eloszláson kívül az általánosított lineáris modellben a függő változó el-
oszlására az exponenciális eloszlások² családjába tartozó binomiális és Poisson-eloszlás
feltételezése a leggyakoribb.

A függő változó várható értéke, μ_i gyakran a magyarázó változók $g(\cdot)$ függvénye, és a
kapcsolatot leíró függvényt a szakirodalom „link” függvénynek nevezi. Így a GLM-
modellt a link függvénnyel felírva:

$$Y_i = \mu_i + \varepsilon_i, \quad \text{ahol } E(Y_i) = \mu_i \quad \text{és} \quad g(\mu_i) = \sum_j x_{ij} \beta_j \quad (i = 1, \dots, n) \quad /2/$$

¹ A GLM további speciális esete például a keresztábrára illesztett loglineáris modell.

² Az exponenciális eloszlás-család magában foglalja a normális, a binomiális és Poisson-eloszlás mellett a gamma és a béta eloszlást is.

adódik, ahol a $g(\cdot)$ jobb oldala a magyarázó változók lineáris prediktora.³

A GLM általános definíciója szerint a modell függő változói az Y_1, Y_2, \dots, Y_n független, az exponenciáliseloszlás-családhoz tartozó, azonos eloszlású valószínűségi változók. Az x magyarázó változók között szerepelhetnek olyan mennyiségi változók, mint az életkor vagy a jövedelem, valamint olyan faktorok – több kategóriát tartalmazó minőségi változók –, mint a nemek, foglalkozások vagy a gépjármű-kockázati csoportok.

EXPONENCIÁLIS ELOSZLÁSOK CSALÁDJÁ

Legyen Y olyan valószínűségi változó, amelynek eloszlása a θ és a ϕ paraméter⁴ függvénye. Az Y eloszlása beletartozik az exponenciális eloszlások családjába, ha valószínűségeloszlása (sűrűségfüggvénye) felírható az alábbi formában:

$$f(y; \theta, \phi) = \exp\left\{\frac{y\theta - b(\theta)}{a(\phi)} + c(y, \phi)\right\}, \quad /3/$$

ahol $a(\cdot)$, $b(\cdot)$, és $c(\cdot, \cdot)$ függvények.

A továbbiakban a normális, a binomiális és a Poisson-eloszlásokról megmutatjuk, hogy az exponenciális eloszlások családjába tartoznak.

a) Normális eloszlás μ várható értékkel (és σ^2 varianciával)

$$f(y) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left[-\frac{1}{2\sigma^2}(y-\mu)^2\right] = \exp\left[\frac{y\mu - \mu^2/2}{\sigma^2} - \frac{1}{2}\left\{\frac{y^2}{\sigma^2} + \log(2\pi\sigma^2)\right\}\right]. \quad /4/$$

A normális eloszlás sűrűségfüggvénye a /3/ szerinti alakra hozható, ha

$$\theta = \mu, \quad b(\theta) = \theta^2/2 = \mu^2/2, \quad a(\phi) = \phi = \sigma^2 \quad \text{és} \quad c(y, \phi) = -(1/2)\{y^2/\phi + \log(2\pi\phi)\}.$$

b) Binomiális eloszlás (n, p) paraméterekkel, n ismert.

Tegyük fel, hogy $Z \sim B(n, p)$. Mivel kényelmesebb a modellek felírása az arányokkal, mint a kedvező esetek számával, ezért legyen $Y = Z/n$.

$$\begin{aligned} f(z) &= \binom{n}{z} p^z (1-p)^{n-z} \quad \text{így} \quad f(y) = P(Y=y) = P(Z=ny) = \binom{n}{ny} p^{ny} (1-p)^{n-ny} = \\ &= \exp\left\{ny \log p + (n-ny) \log(1-p) + \log\left(\binom{n}{ny}\right)\right\} = \exp\left\{ny \log \frac{p}{1-p} + n \log(1-p) + \log\left(\binom{n}{ny}\right)\right\} = \\ &= \exp\left[n\left\{y \log \frac{p}{1-p} + \log(1-p)\right\} + \log\left(\binom{n}{ny}\right)\right]. \quad /5/ \end{aligned}$$

³ A függő változóra felírt GLM alternatívát jelent a magyarázó változók transzformációjával szemben.

⁴ Az egyszerűbb tárgyalás kedvéért tegyük fel, hogy a ϕ értéke ismert.

Ez a modell a /3/ szerinti alakot ölti, ha

$$\theta = \log \frac{p}{1-p} \text{ és így } p = \frac{e^\theta}{1+e^\theta},$$

továbbá

$$b(\theta) = \log(1 + \exp(\theta)) = -\log(1 - p), \phi = n, a(\phi) = 1/\phi \text{ és } c(y, \phi) = \log \binom{n}{ny}.$$

c) Poisson-eloszlás λ várható értékkel

$$f(y) = \exp(-\lambda) \frac{\lambda^y}{y!} = \exp\{y \log \lambda - \lambda - \log y!\}. \quad /6/$$

Ez a függvény megfelel a /3/ szerinti képletnek, ha

$$\theta = \log \lambda, b(\theta) = \exp(\theta) = \lambda, a(\phi) = \phi = 1 \text{ és } c(y, \phi) = -\log y!$$

AZ EXPONENCIÁLIS ELOSZLÁSCSALÁD Y VÁLTOZÓJÁNAK VÁRHATÓ ÉRTÉKE ÉS VARIANCIÁJA

Az Y változó várható értéke és varianciája a $b(\theta)$ és az $a(\phi)$ függvények segítségével határozható meg. A likelihood függvény legyen:

$$L = \prod_{i=1}^n f(y_i, \theta, \phi)$$

Jelölje a log-likelihood függvényt $\ell = \log L$, és a score⁵ függvényt $U = d\ell/d\theta = d \log L / d\theta$. A score függvény alábbi két tulajdonságát használjuk fel Y várható értékének és szórásnégyzetének előállításához:

$$E(U) = E(d \log L / d\theta) = 0, \quad /7/$$

$$\text{Var}(U) = -E(dU/d\theta) = -E(d^2 \log L / d\theta^2). \quad /8/$$

Az exponenciáliseloszlás-család logaritmusát /3/ alapján

$$\ell = \{y\theta - b(\theta)\} / a(\phi) + c(y, \phi)$$

⁵ A „score” itt a log-likelihood függvénynek a vizsgált paraméter szerinti parciális deriváltja. Nincs elfogadott magyar neve.

és így

$$U = d\mathcal{L}/d\theta = \{y - b'(\theta)\}/a(\phi). \quad /9/$$

A /7/ alapján $E(U) = 0$ miatt $E(Y) = b'(\theta)$.

A /8/ szerinti deriváltra $-dU/d\theta = b''(\theta)/a(\phi)$ adódik, ezért a /9/ szerinti U függvény varianciája:

$$\text{Var}(U) = E(b''(\theta)/a(\phi)) = \text{Var}(Y)/a^2(\phi). \quad /10/$$

Fejezzük ki /10/-ből $\text{Var}(Y)$ -t:

$$\text{Var}(Y) = a^2(\phi)E(b''(\theta)/a(\phi)) = a(\phi)b''(\theta). \quad /11/$$

A $b''(\theta)$ -t varianciafüggvénynek nevezi és $\text{Var}(\mu)$ -vel jelöli a szakirodalom, mert azt mutatja meg, hogyan függ y_i varianciája az y_i várható értékétől.

Vizsgáljuk meg a varianciafüggvény alakját a három nevezetes eloszlás esetében.

1. Normális eloszlást felírva $b(\theta) = \theta^2/2$, és így $\mu = E(Y) = b'(\theta) = \theta$, a szórásnégyzet pedig $\text{Var}(Y) = a(\phi)b''(\theta) = \phi \cdot 1 = \sigma^2$. Mivel $b''(\theta) = 1$, a varianciafüggvény $\text{Var}(\mu) = 1$, azaz konstans.⁶

2. Binomiális eloszlást feltételezve $b(\theta) = \log(1 + \exp(\theta))$ és így $\mu = E(Y) = b'(\theta) = \exp(\theta)/\{1 + \exp(\theta)\} = p$, továbbá $\text{Var}(Y) = a(\phi)b''(\theta) = (1/n)[\exp(\theta)/\{1 + \exp(\theta)\}^2] = p(1-p)/n$. A második deriváltra $b''(\theta) = p(1-p) = \mu(1-\mu)$ adódik, ezért a varianciafüggvény $\text{Var}(\mu) = \mu(1-\mu)$.

3. Poisson-eloszlást feltételezve $b(\theta) = \exp(\theta)$, és ezért $\mu = E(Y) = b'(\theta) = \exp(\theta) = \lambda$, továbbá a szórásnégyzet is megegyezik a várható értékkel: $\text{Var}(Y) = a(\phi)b''(\theta) = \exp(\theta) \cdot 1 = \lambda = \mu$. Mivel $b''(\theta) = \mu$, a varianciafüggvény: $\text{Var}(\mu) = \mu$.

AZ EXPONENCIÁLIS ELOSZLÁSCSALÁD HÁROM TAGJÁRA ILLESZTHETŐ ÁLTALÁNOSÍTOTT LINEÁRIS MODELLEK

Ha a $g(\cdot)$ függvény éppen megegyezik az exponenciális eloszláscsaládból választott eloszlás paraméterét és a θ -t összekapcsoló függvénnyel, akkor a kapcsolatot leíró függvényt *kanonikus link*-nek nevezzük.

a) Normális eloszlású Y változó, a várható értéke: μ .

A /4/-ből leolvasható, hogy $\theta = \mu$, ezért az azonosság teremti meg a kanonikus kapcsolatot:

$$g(\mu) = \mu. \quad /12/$$

⁶ A normális eloszlás konstans varianciafüggvénye összhangban van azzal a nevezetes tulajdonsággal, hogy a normális eloszlás két paramétere (a várható érték és a variancia) független.

b) Binomiális változó esetén $Z \sim B(n,p)$ az arány $Y = Z/n$.

Az /5/-ből következően $\theta = \log \frac{p}{1-p} = \log \frac{\mu}{1-\mu}$, és így az esélyhányados logaritmus, a logit a kanonikus függvény:

$$g(\mu) = \log \frac{\mu}{1-\mu} . \tag{13/}$$

c) Poisson-eloszlás esetén $Y \sim \text{Poisson}(\lambda)$.

A /6/ alapján $\theta = \log \lambda = \log \mu$, és így látható, hogy a kanonikus kapcsolatot a logaritmus függvény teremti meg:

$$g(\mu) = \log \mu . \tag{14/}$$

Eddigi eredményeinket az 1. tábla foglalja össze. Ha a θ és a $g(\cdot)$ oszlopok tartalmát összevetjük, akkor egyértelművé válik, hogy a θ -t azért nevezik kanonikus paraméternek, mert megegyezik a kanonikus kapcsolatot leíró $g(\cdot)$ függvénnyel.

1. tábla

Az exponenciális eloszláscsalád paramétereinek közötti kapcsolat

Eloszlás	$E(Y) = \mu$	θ	ϕ	$\text{Var}(\mu)$	$g(\cdot)$ link
Normális (μ, σ^2)	μ	μ	σ^2	1	μ
Binomiális (n, p)	np	$\log p/(1-p)$	n	$p(1-p)/n$	$\log p/(1-p)$
Poisson (μ)	μ	$\log \mu$	1	μ	$\log \mu$

AZ ÁLTALÁNOSÍTOTT LINEÁRIS MODELL PARAMÉTEREINEK MAXIMUM LIKELIHOOD BECSLÉSE

A GLM-modell alkalmazása során ismertnek tételezzük fel a függő változó valószínűségeloszlását, ezért a likelihood függvény úgy írható fel, mint az exponenciális eloszláscsalád sűrűségfüggvényeinek szorzata. A log-likelihoodot maximalizáljuk, hogy meghatározzuk a lineáris prediktorban szereplő β paraméterek értékét, majd meghatározzuk a becslés standard hibáját.

A GLM-modellek adatokhoz történő illesztése számítógépes programcsomagokkal, például az Splus-szal vagy az SPSS-szel végezhető el. Ha feltáró, vagy – új szóhasználattal – adatvezérelt elemzést végzünk, akkor általában úgy járunk el, hogy a modellváltozatok egész sorát illesztjük az adatokhoz. Az egyes lépésekben bizonyos változókat és faktorokat kizárunk, illetve bevonunk a modellbe, hogy mérni tudjuk kizárásuk, illetve bevonásuk hatását a modell illeszkedésére.

A GLM-modellnek a megfigyelt adatokhoz való illeszkedését jellemezhetjük:

- a modellváltozatok között a log-likelihood függvények eltérésének mérőszáma (*deviancia*) alapján,
- a becsült és a megfigyelt értékek közti reziduumok vizsgálatával.

Modell-deviancia

Tegyük fel, hogy $p(< n)$ számú paramétert tartalmazó általánosított lineáris modellt illesztünk az y_1, \dots, y_n megfigyelt adatokra. A maximum likelihood (ML) becsléssel kapott paraméterekkel felírt log-likelihood értékét jelölje $\ell(\hat{\beta})$.

Ha ugyanilyen (azonos eloszlású függő változót és link függvényt tartalmazó) modellt illesztünk megfigyeléseinkre, de a modell most n darab, azaz a megfigyelésekkel megegyező számú paramétert tartalmaz, akkor telített modellt kapunk.

A telített modell tökéletesen illeszkedik az adatokhoz, ezért az ML becsléssel kapott paramétervektor mellett a log-likelihood felveszi a maximumát: $\ell(\hat{\beta}_{\max})$. Az első modell jól leírja az adatokat, ha $\ell(\hat{\beta})$ közel van a telített modell log-likelihood értékéhez, $\ell(\hat{\beta}_{\max})$ -hoz; és gyenge a modell illeszkedése, ha $\ell(\hat{\beta})$ sokkal kisebb, mint $\ell(\hat{\beta}_{\max})$.

A modellek log-likelihoodja közötti különbség kétszeresét, a devianciát jelölje D :

$$D = 2[\ell(\hat{\beta}_{\max}) - \ell(\hat{\beta})]. \quad /15/$$

Ha a minta elég nagy, akkor a devianciát mérő statisztika közelítőleg χ^2 eloszlású, szabadságfoka a paraméterek számában elért csökkenés: $(n-p)$.

$$D \sim \chi_{n-p}^2$$

A D mutatót akkor is kiszámíthatjuk, ha a modell csak a konstans tagot tartalmazza, ekkor a jele D_0 , és neve null-deviancia, szabadságfoka pedig $(n-1)$.

Tegyük fel, hogy az M_2 modellnek p darab paramétere és D_2 devianciája,⁷ míg az M_1 modellnek q számú paramétere ($q < p$) és D_1 devianciája van, vagyis az M_1 modell beágyazott modellje (speciális esete) az M_2 modellnek. Felmerül a kérdés, hogy a $(p-q)$ extra paraméter bevonása az M_2 modellbe lényegesen jobb illeszkedést biztosít-e, mint az egyszerűbb M_1 modell. Az illeszkedés javulását a deviancia változásával mérhetjük, mert $D_1 - D_2$ is χ_{p-q}^2 eloszlást követ. Ha a két modell egyformán hatékony, akkor a $D_1 - D_2$ értéke kicsi; ha pedig M_2 lényegesen jobban illeszkedik, mint M_1 , akkor $D_1 - D_2$ viszonylag nagy. Az M_1 és M_2 közötti választásra jobb oldali kritikus tartománnyal rendelkező χ^2 -próbát használunk.⁸ Hüvelykujj-szabályként olykor használhatjuk azt, hogy az M_2 modellt preferáljuk M_1 ellenében, ha $D_1 - D_2 > 2(p - q)$.

Reziduumok

Az előzőekben láttuk, hogy a deviancia alapján összehasonlíthatjuk különböző modellek illeszkedésének jóságát. Az egyes modellek további vizsgálatához felhasználhatjuk a megfigyelt és az illesztett értékek eltérését, azaz a reziduumot, amely az i -edik megfi

⁷ A modellváltozatok reziduális devianciái közvetlenül is összehasonlíthatók, mert ha két modell D mutatójának különbségét vesszük, akkor a /15/-ből éppen kiesik a telített modell log-likelihoodja.

⁸ Ez a próba megegyezik a likelihood arány vagy más néven maximum likelihood arány teszttel.

gyelésre $y_i - \hat{\mu}_i$ értékű. A reziduumok összege zérus, és az egyes eltérések jelentősége sem ítéhető meg a megfigyelt vagy az illesztett érték nagyságának ismerete nélkül. Ezért a reziduumokat általában standardizálás után vizsgáljuk.

Az egyes megfigyelések hatását a *reziduum-deviancia*⁹ értéke alapján minősítjük, amelynek előjele megegyezik az $(y_i - \hat{\mu}_i)$ egyedi eltérés előjével, nagysága pedig $\sqrt{D_i}$. Így a reziduum-devianciák négyzetösszege éppen a /15/ szerinti D devianciát, azaz a likelihood arány teszt próbafüggvényének értékét adja.

ÁLTALÁNOSÍTOTT LINEÁRIS MODELLEK BIZTOSÍTÁSI ALKALMAZÁSAI

1. Halandósági becslések. Példaként tételezzük fel, hogy Y_x egy adott évben a legközelebbi születésnapjuk szerint x éves korú népességben bekövetkezett halálesetek számát fejezi ki. Az x évesek kockázati időtartamát¹⁰ jelölje E_x^c , és a μ_x halálozási intenzitás Gompertz-törvénye¹¹ szerint alakul:

$$\mu_x = Bc^x.$$

Ekkor egy lehetséges modell az alábbi alakban írható fel:

$$Y_x \sim \text{Poisson}(E_x^c \mu_x), \text{ ahol } \mu_x = Bc^x. \quad /16/$$

A Poisson-eloszlású változó várható értéke az eloszlás paramétere, ezért

$$E(Y_x) = E_x^c \mu_x \quad /17/$$

és így $\log E(Y_x) = \log E_x^c + \log Bc^x = \log E_x^c + \log B + (\log c)x$, azaz

$$\log E(Y_x) = a + bx. \quad /18/$$

Itt tehát az Y függő változó várható értékének logaritmusa az életkor lineáris függvénye, azaz a két változó közti kapcsolatot a logaritmusfüggvény teremti meg, ezért itt a kanonikus kapcsolatot alkalmazzuk. A lineáris prediktor $g(\mu_i) = a + bx_i$, ahol $a = \log B$ és $b = \log c$.

Halandósági modellünkben definiáljuk a Z halálozási rátát az x éves korban meghaltak számának és a kockázati időtartamnak az arányával: $Z_x = Y_x/E_x$. Ekkor egy adott x_i életkorban¹² megfigyelünk Y_i számú halálesetet az E_i kockázatban töltött időre, és ezek arányaként kapjuk a Z_i halálozási rátát, ahol $Z_i = \mu_i + \varepsilon_i$. Legyen a halálozási ráta várható

⁹ A reziduum-devianciák aszimptotikusan normális eloszlást követnek.

¹⁰ Angolul az *exposure* vagy az *exposed to risk* kifejezés szerepel. Ez a biztosított kockázatnak „kitettséget”, azaz azt az időtartamot fejezi ki, amíg az adott évben a biztosított az állományba tartozik.

¹¹ Gompertz halandósági függvénye 1825-ből való. A függvényben szereplő paraméterekre az alábbi megkötések érvényesek: $B > 0$, $c > 1$ és $x \geq 0$. A Gompertz-függvény tulajdonságait részletesen lásd *Valkovics*; 2000.

¹² A továbbiakban egyszerűsítjük a jelölést, és Y_x helyett Y_i jelöli az $x = x_i$ életkorban meghaltak számát.

értéke /16/ szerint a halálozási intenzitás: $E(Z_i) = \mu_i$, ezért a /18/-ból $\log \mu_i = a + bx_i$ adódik, ahol $a = \log B$ és $b = \log c$. Tehát a halálozási rátát is a logaritmusfüggvény, mint kanonikus link becsli.

2. *Gépjárműkár-adatok becslése.* A biztosítónál rendelkezésre áll a kötvénytulajdonosok néhány adata, többek között a nemük, az életkoruk, a lakóhelyük, a gépjármű kockázati besorolása, és az adott időszak alatt bekövetkezett károk száma.

Tegyük fel, hogy az i -edik kötvénytulajdonosnál bekövetkezett károk száma λ_i paraméterű Poisson-eloszlást követ, ahol a λ_i logaritmus a kockázati faktorok lineáris függvénye :

$$\log \lambda_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}.$$

A modell felépítése során különböző megfontolásokat követhetünk. Kiválaszthatunk egyetlen fontosnak tartott változót, bevonhatunk további változókat és faktorokat, valamint szerepeltethetjük ezek interakcióit is az alábbiak szerint.

1. modell: életkorváltozó.

Ha csak az x_1 -gyel jelölt életkorváltozót (értékei: 17, 18, 19, ...) szerepeltetjük a lineáris prediktorban, akkor $\beta_0 + \beta_1 x_1$ becsli a károk várható számának logaritmusát.

2. modell: életkor + nem.

Az életkor mellett a nemek szerinti különbség is döntően befolyásolhatja a kárszámok alakulását. A „nem” olyan faktor, amely két lehetséges értéket vesz fel: férfi = 0 és nő = 1. Az ilyen bináris változó többféleképpen is beépíthető a modellbe. Ha a változó értéke $x_2=0$ a férfiak és 1 a nők esetében, akkor a prediktor $\beta_0 + \beta_1 x_1 + \beta_2 x_2$, vagy egyszerűbben felírva: $\beta_0 + \beta_1 x_1 + \alpha$ (ahol α jelenti a károk számában a „női” hatást).

3. modell: életkor + nem + kockázati csoport.

A biztosító a gépjárműveket például három kockázati csoportba sorolja, így ez a faktor háromfokú ordinális skálán mér. Ha az első csoport jelenti az alapszintű kockázatot, akkor a további két csoport extra kockázatát a γ_1 és γ_2 paraméterek fejezik ki, és a modellben két magyarázó „dummy” változó szerepel. Lineáris prediktorunk az $(i + 1)$ -edik kockázati csoportban a következő: $\beta_0 + \beta_1 x_1 + \alpha + \gamma_i$. ($i = 1, 2$)

4. modell: életkor + nem + kockázati csoport + (nem) · (kockázati csoport).

A GLM-modell fontos kiterjesztését jelenti az a lehetőség, hogy figyelembe vehető az egyes változók és faktorok interakciója. Példaként említhető az, hogy a gépjárművek kockázati besorolása másként hat a kárszámra, ha fiatal vagy ha tapasztalt vezető ül egy azonos kockázati csoportba sorolt autó volánjánál. Az életkor és a kockázati csoport közötti interakció hatását a modellben egy újabb paraméter fejezi ki:

$$\beta_0 + \beta_1 x_1 + \alpha + \gamma_i + \delta_i.$$

A számítások elvégzéséhez a (feltételezett) biztosítási állományból válasszuk ki véletlenszerűen 55 kötvénytulajdonos adatait. A függő változó legyen az egy év alatt egy

kötvényen bekövetkező károk száma. A kárszám a feltételezések szerint Poisson-eloszlást követ, $Y_i \sim \text{Poisson}(\lambda_i)$, ezért a (kanonikus) log link függvényt írjuk fel például a 3. modell szerint:

$$\log \lambda_i = \beta_0 + \beta_1 x_i + \alpha + \gamma_i,$$

ahol β_0 a konstans tag, β_1 az életkori hatás együtthatója (x az életkor változó), α a női vezetők kárhatását méri, és γ_1 valamint γ_2 a B és C kockázati csoportba¹³ sorolt gépjárművek káralakulásra gyakorolt többlethatását fejezi ki (ha az A kockázati csoport jelenti az alapkockázatot). Mivel az A kockázatu járművet vezető férfiakra írjuk fel a legegyszerűbb modellt ($\beta_0 + \beta_1 x_i$), ők jelentik a kárszám becslésénél a referenciacsoportot.

Hét GLM-modellt illesztettünk az Splus szoftverrel, és a számítások a következő eredményekre vezettek:

- a magyarzó változót nem tartalmazó null-modell devianciája $D_0 = 66,306$, szabadságfoka 54;
- az egyváltozós modellek deviancia-mértéke jelentősen különbözik egymástól.

Az M_1 modellben az életkor változó bevonásával a $D_1 = 61,146$, szabadságfoka 53. A deviancia csökkenése M_0 -hoz viszonyítva 5,160 és $D_0 - D_1$ szabadságfoka egy.

Ha csak a nem, mint magyarzó változó épült be az M_2 modellbe, akkor a deviancia értéke 66,301 (szabadságfoka 53), és ez csak 0,005-es csökkenést jelent. A „nem” változó bevonása tehát nem javította szignifikánsan a modell illeszkedését.

Amikor a kockázati csoport volt az M_3 modellben az egyetlen magyarzó változó, akkor a $D_3 = 56,168$ (szabadságfoka 52), és ez a null modell devianciáját 10,138-del csökkenteti 2 szabadságfok mellett.

Az egyváltozós modellek közül M_1 és M_3 preferálható M_0 -val szemben. Feltételezhető, hogy az illeszkedés tovább javítható újabb magyarzó változó vagy faktor bevonásával.

A fenti eredmények alapján a kétváltozós modellek közül csak a kockázati csoportot is tartalmazó M_4 és M_5 változatokat számszerűsítettük.

A kockázati csoport és a nemek együttes szerepeltetése az M_4 -ben 56,136 devianciát eredményez (51 szabadságfokkal). A férfi–nő megkülönböztetés a modell illeszkedését nem javítja szignifikánsan, mert 1 szabadságfok mellett $D_3 - D_4 = 0,032$.

Az ötödik modellben a kockázati csoport és az életkor prediktora szerepel, az eltérések mértéke $D_5 = 50,208$ (a szabadságfok itt is 51). Az M_3 modellhez képest a javulás szignifikáns, mert $D_3 - D_5 = 5,95$ és ez jóval nagyobb, mint az egységnyi szabadságfok kétszerese.

A hatodik modell mindhárom magyarzó változót tartalmazza, és így a deviancia 50,205 és szabadságfoka 50. Az illeszkedés javulása nem szignifikáns, mert $D_5 - D_6 = 0,003$ és a szabadságfok 1.

Az eddigi eredmények alapján az ötödik modellt választjuk, melynek számítógépes eredményeit¹⁴ a 2. tábla mutatja. A magyarzó változók a táblában: a kötvénytulajdonos életkora és a kockázati csoport.

¹³ Az A jelölje a kicsi, olcsó és nem túl erős motorú kocsikat, a B a közepes tulajdonságúakat, és a C kategóriába tartoznak az erős, gyors és drága autók.

¹⁴ A konstans a szokásos szignifikanciaszintek mellett nem különbözik zérustól. A t -teszt kritikus értéke 50 szabadságfok és 5 százalékos kétoldali szignifikanciaszint mellett 2,01.

2. tábla

Kockázati csoport és életkor-együtthatók becslési értéke a GLM-ben

Megnevezés	Paraméter	Standard hiba	t-teszt
(Konstans)	-0,17138	0,5432	-0,3155
B kockázat	1,09783	0,4286	2,5614
C kockázat	1,51887	0,5166	2,9402
Életkor	-0,02922	0,0130	-2,2447

Null-deviancia: 66,306 Szabadságfok: 54
 Reziduális deviancia: 50,208 Szabadságfok: 51

E modell alapján a lineáris prediktorok számítása néhány személyre a következő:

– 36 éves (férfi) A kockázati csoportba sorolt autóval:

$$\beta_0 + 36\beta_1 = -0,17138 + [36 \times (-0,02922)] = -1,2233;$$

– 24 éves (nő) C kockázati autóval (α nem szerepel a modellben):

$$\beta_0 + 24\beta_1 + \gamma_2 = -0,17138 + [24 \times (-0,02922)] + 1,51887 = 0,6463;$$

– 53 éves (férfi) B kockázati gépjárművel:

$$\beta_0 + 53\beta_1 + \gamma_1 = -0,17138 + [53 \times (-0,02922)] + 1,09783 = -0,6222.$$

A Poisson-eloszlást feltételező modellben $\log\lambda$ -t becsljük, így λ becslési értéke $\exp(\beta_0 + \beta_1 \text{életkor} + \text{kockázat})$, és ezért az illesztett várható kárszámok az előbb bemutatott három személy esetében a következők:

- 36 éves, alapkockázati vezető becslési kárszáma: $\exp(-1,2233) = 0,294$;
- 53 éves, B kockázati csoportba sorolt vezető becslési kárszáma: $\exp(-0,6222) = 0,537$;
- 24 éves, C csoportba sorolt vezető becslési kárszáma: $\exp(0,6463) = 1,908$.

A hetedik lépésben az ötödik modell illeszkedését a kockázati csoport és az életkor kölcsönhatásának szerepeltetésével kívántuk javítani.

3. tábla

Interakciót is tartalmazó GLM illesztése

Megnevezés	Paraméter	Standard hiba	t-teszt
(Konstans)	-0,0376	0,8342	-0,0451
B kockázat	0,6287	1,0677	0,5889
C kockázat	2,4646	1,5479	1,5923
Életkor	-0,0335	0,0245	-1,3676
KorxB kockázat	0,0132	0,0293	0,4484
KorxC kockázat	-0,0307	0,0492	-0,6243

Null-deviancia: 66,306 Szabadságfok: 54
 Reziduális deviancia: 49,106 Szabadságfok: 49

A deviancia mértéke $55 - 6 = 49$ szabadságfok mellett 49,106, és $D_5 - D_7 = 1,102$, ami 1 szabadságfok mellett nem jelent szignifikáns csökkenést. A számítógépes eredményekből (3. tábla) látható, hogy az interakció együtthatója egyik kockázati csoportban sem különbözik szignifikánsan zérustól, sőt maguk a kockázati csoportok is elveszítették statisztikai jelentőségüket.

A legjobb modellnek tehát az életkort és kockázati csoportokat mint magyarázó változókat tartalmazó 5. modell tekinthető. A 4. táblában az 5. modellel becsült kárszámok szerepelnek.

4. tábla

Életkor és kockázati csoport alapján becsült kárszámok

Életkor (év)	A kockázati csoport	B kockázati csoport	C kockázati csoport
17	0,513	1,537	2,342
24	0,418	1,252	1,908
36	0,294	0,882	1,344
53	0,179	0,537	0,818

A díjkalkuláció során Magyarországon is használják a GLM-becslést, amellyel például megállapítható, hogy a 17 éves A kockázati vezető és az 53 éves B típusú vezető autós közel azonos díjat fizethet, mert várt kárgyakoriságuk nagyon hasonló.

HALÁLOZÁSI ARÁNY BECSLÉSE LOGISZTIKUS REGRESSZIÓVAL

A GLM-modellt arányra is felírhatjuk. Legyen Y_i az n_i elemű csoportban a vizsgált esetek arányát mérő változó, és így $\mu_i = E(Y_i) = p_i$.

Ha halálozási adatokhoz illesztünk általánosított lineáris modellt, akkor binomiális eloszlást tételezünk fel, és a /13/ alapján: $Y_i = p_i + \varepsilon_i$, ahol $\log[p_i / (1 - p_i)] = a + bx_i$.

A kapcsolatot leíró függvényben az esélyhányados logaritmus $g(p) = \log[p / (1 - p)]$ szerepel, ezért ezt a modellt a szakirodalom az arányokra felírt logisztikus regressziós modellnek, vagy röviden *logit modellnek* nevezi.

Hasonlóan jól alkalmazható a logit modell különböző kezelések, hatóanyagok, például rovarirtó szerek hatásainak¹⁵ tesztelésére. Tegyük fel, hogy a kísérleti egyedeket k számú csoportba osztjuk, és az i -edik csoportban n_i számú egyed van. Az egyes csoportok minden tagja x_i dózist kap az adott hatóanyagból, és az i -edik csoportban z_i számú pusztulás következik be (és $n_i - z_i$ egyed túléli a kezelést).

Becsülni kívánjuk azt a p_i valószínűséget, hogy egy egyed elpusztul, ha x_i mennyiséget kap a hatóanyagból. Ha Z jelöli az n elemű csoportban az elpusztultak számát, akkor $Z(n, p)$ paraméterű binomiális eloszlást követ, nem pedig normális eloszlást. A p valószínűség és az x magyarázó változó között a kapcsolat általában nem lineáris. Mivel a becslés során a bevezetőben említett mindkét probléma felmerül, célszerű az általánosított lineáris modellt használni.

¹⁵ Ezért az angol nyelvű szakirodalomban a GLM-modell függő változóját „response”-ként és nem „dependent”-ként említik.

Mintapéldánkban, ahol egy rovarirtó hatását elemezzük 10 csoportba osztva vizsgálunk 607 egyed. Az egyes csoportok különböző dózisu kezelést/vegyszert kaptak, ezt a magyarázó változó (dózis) méri (a megfelelő mértékegységben). A függő változó az elpusztult rovarok arányát adja meg csoportonkénti bontásban.

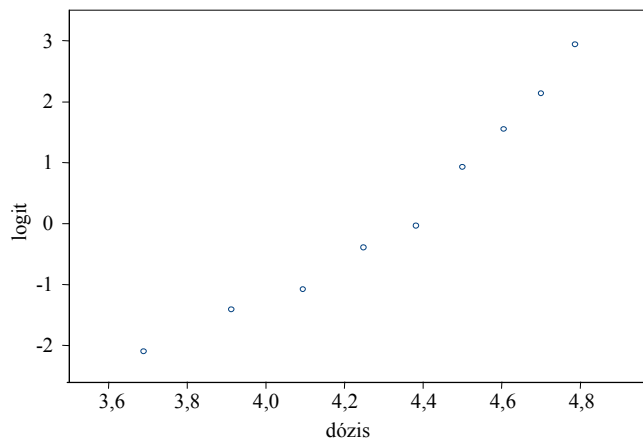
5. tábla

Csoportonkénti pusztulási adatok és dózisok

Csoport	A csoport mérete	Az elpusztult rovarok száma	Az elpusztult rovarok aránya	Dózis
1	64	7	0,1094	3,689
2	66	13	0,1970	3,912
3	59	15	0,2542	4,094
4	62	25	0,4032	4,248
5	55	27	0,4909	4,382
6	60	43	0,7167	4,500
7	63	52	0,8254	4,605
8	57	51	0,8947	4,700
9	60	57	0,9500	4,787
10	61	61	1,0000	4,868

Az 1. ábra vízszintes tengelyén a dózis, függőleges tengelyén a logit alakulása látható.

1. ábra. Dózis és pusztulás kapcsolata



Szembevető az 1. ábrán, hogy a két változó kapcsolata nem lineáris, tehát várható, hogy a lineáris prediktor illesztésével nem kapunk megfelelő modellt. A logisztikus regresszió alkalmazásával a null-modell devianciája 298,067. A 9-es szabadságfoknál 5 százalékos szignifikanciaszinten a χ^2 -eloszlás kritikus értéke 16,679, tehát ez a modell nem megfelelő.

Ha a dózist mint magyarázó változót bevonjuk a modellbe, akkor a 8 szabadságfokú modell devianciája 17,234, ami a null-modellhez képest szignifikáns csökkenést jelent 1 szabadságfok mellett. A számítások eredményét a 6. tábla tartalmazza.

6. tábla

Lineáris prediktor illesztése logit modellel

Megnevezés	Paraméter	Standard hiba	t-teszt
(Konstans)	-20,2723	1,5925	-12,7297
Dózis	4,7331	0,3654	12,9535

Null-deviancia: 298,0666 Szabadságfok: 9
 Reziduális deviancia: 17,23397 Szabadságfok: 8

A prediktor értéke 4 egységnyi dózis esetén: $-20,2723 + (4 \times 4,7331) = -1,3399$. A logit kapcsolat miatt az elpusztult rovarok arányának illesztett értéke

$$p = \frac{\exp(\beta_0 + \beta_1 x)}{1 + \exp(\beta_0 + \beta_1 x)},$$

azaz

$$\exp(-1,3399)/[1 + \exp(-1,3399)] = 0,208.$$

Természetes, hogy a dózis mint magyarázó változó szerepeltetése a null-modellhez képest jelentősen javítja a modell illeszkedését. A megfigyelt adatokat jól leíró modell közelítően χ^2_8 eloszlású. A deviancia értéke (17,234) azonban nagyobb, mint a táblabeli kritikus érték, amely 5 százalékos valószínűségi szinten 15,507. A modell illeszkedése tökéletesíthető, ha a prediktorban kvadratikusan tag is szerepel:

$$\log \frac{p_i}{1 - p_i} = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i}, \text{ ahol } x_{2i} = x_{1i}^2.$$

Az illesztés eredményét a 7. tábla tartalmazza.

7. tábla

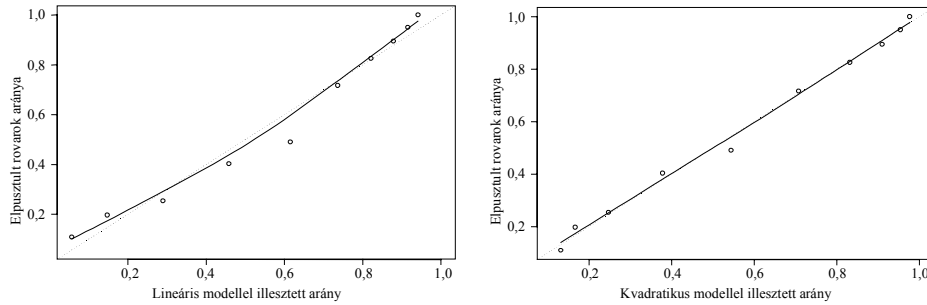
Kvadratikusan prediktor illesztése logit modellel

Megnevezés	Paraméter	Standard hiba	t-teszt
(Konstans)	47,1539	18,8194	2,5056
Dózis	-26,9889	8,9119	-3,0284
Dózis négyzete	3,7126	1,05133	3,5313

Null-deviancia: 298,0666 Szabadságfok: 9
 Reziduális deviancia: 4,590 Szabadságfok: 7

A deviancia 4,590, a szabadságfok 7, és ez 12,644-del csökkenti a lineáris modell devianciáját. Négy egységnyi dózis mellett a prediktor $47,1539 + [4 \times (-26,9889)] + (4^2 \times 3,7126) = -1,4000$, és a becsült pusztulási arány $\exp(-1,4)/[1 + \exp(-1,4)] = 0,198$.

2. ábra. A lineáris és a kvadratikus prediktor illesztése



A kvadratikus modellel készített becslés lényegesen jobban illeszkedik az adatokhoz mind a modell devianciaértéke, mind a 2. ábra jobb oldali része alapján. Az ábrák vízszintes tengelyén a logit modellel illesztett, a függőleges tengelyen pedig a megfigyelt pusztulási arány (5. tábla) szerepel.

IRODALOM

- DOBSON, A. J. (1990): *An introduction to Generalized Linear Models*. Chapman&Hall, 174 old.
- FÜSTÖS, L. – KOVÁCS, E. (1989): *A számítógépes adatelemzés statisztikai módszerei*. Tankönyvkiadó, Budapest. 380 old.
- JOBSON, J. D. (1992): *Applied multivariate data analysis*. Volume II. Springer-Verlag, 731 old.
- KRZANOWSKI, W. J. (1998): *An introduction to statistical modelling*. Arnold, 252 old.
- McCULLOCH, C. E. – SEARLE, S. R. (2001): *Generalized, linear, and mixed models*. Wiley Series in Probability and Statistics, 325 old.
- VALKOVICS, E. (2001): A Gompertz-függvény felhasználási lehetőségei a demográfiai modellezésben. *Statisztikai Szemle*, 79. évf. 2.sz. 121–141. old.
- VENABLES, W. N. – RIPLEY, B. D. (1999): *Modern applied statistics with S-Plus*. Springer-Verlag 3rd ed. 520 old.

SUMMARY

There are two main reasons why the linear regression model may be unsatisfactory in some applications. The response variable may not be normally distributed, or the mean of the response variable is a function of the linear predictor, rather than the predictor itself.

The GLM is used to fit a model under different distributional assumptions. Logit is the link function for the binomial distribution, logarithm link is used for the Poisson distribution, and the identity link is fitted in case of the Gauss distribution. The explanatory (predictor) variables can be mixed, quantitative variables, or factors with a number of different categories. The goodness-of-fit is measured by the likelihood ratio or scaled deviance, which statistics follow chi-square distribution.

The generalized linear model can be useful for insurance data including linear predictor for the response variable. Gompertz law of mortality to Y_x , the number of deaths at age x is an example of fitting GLM. Car rating groups are estimated under Poisson distribution assumption, when logit link is the canonical link. Several models are fitted, and tested using the deviance measure.

A CHAID ALAPÚ DÖNTÉSI FÁK JELLEMZŐI

HÁMORI GÁBOR

Napjainkban, elsősorban a fejlett statisztikai kultúrájú országokban, egyre szélesedik a döntési, más néven klasszifikációs fák alkalmazási köre. Segítségükkel döntési szabályok hozhatók létre, szegmentálásra, osztályozásra nyílik lehetőség. A fák működésének háttérben meglehetősen bonyolult statisztikai algoritmusok állnak, melyek közül napjainkban négy eljárásnak van letisztult módszertana. Ez a CHAID (Chi-squared Automatic Interaction Detector), az Exhaustive CHAID, a C&RT (Classification and Regression Trees) és a QUEST (Quick, Unbiased, Efficient Statistical Tree).

Jelen tanulmány célja, hogy az első két eljárás, a CHAID és az Exhaustive CHAID algoritmus működését bemutassa.

TÁRGYSZÓ: Automatikus osztályozás. Döntési fák.

A CHAID, olyan többváltozós rekurzív klasszifikáló eljárás, melyet *G. Kass* fejlesztett ki 1980-ban. Az eljárást *Kass* eredetileg kategóriás kimenetű változókra fejlesztette ki, de később sor került az algoritmus továbbfejlesztésére, mely így már alkalmassá vált mind a függő változó, mind a független változók esetében folytonos ismérvek kezelésére. Az exploratív algoritmus fő célja, hogy a megfigyeléseket a függő változó (Y) szempontjából úgy csoportosítsuk, hogy a csoportokon belüli variancia minél kisebb, míg a csoportok közötti variancia minél nagyobb legyen. Az eljárás során kirajzolódik a magyarázó változók (X_i) hierarchiája is aszerint, hogy a célváltozó varianciáját mekkora mértékben magyarázzák.

Mindezek miatt a CHAID kedvelt szegmentációs technika is és mint ilyen méltó vetélytársa a hagyományos klaszteranalízisnek, mely alapvetően csak mennyiségi változókkal leírható megfigyelések csoportosítására alkalmas. Gyors elterjedésének és népszerűségének fő oka, hogy a kijelölt függő változó és a magyarázó változók közötti kapcsolatrendszer vizuális formában, könnyen értelmezhető fastruktúrában (decision/classification tree) lehet láttatni. Könnyű interpretálhatósága következtében különös népszerűségnek örvend az adatbányászok körében. A modellkészítő statisztikus szempontjából az eljárás nagy előnye, hogy a változók mérési skálájára és azok eloszlására vonatkozóan semmilyen megkötést nem követel meg, folytonos és kategóriás függő és független változókat egyaránt képes kezelni.

AZ ALAPMODELL

Kezdjük az algoritmussal való ismerkedést először a Kass által kifejlesztett eredeti eljárással, melynél a függő változó és a magyarázó változók egyaránt kategóriásak. A függő változó (Y) kijelölése után a CHAID-modellt alkotó rekurzív algoritmus három fő lépésből áll:

- minden egyes magyarázó változó esetében, a függő változóra vonatkozóan, statisztikailag független, pontosabban a statisztikailag legkevésbé összefüggő kategóriák *egyesítése* (merging);
- a megfigyelések, a függő változó tekintetében legkevésbé függetlennek tekinthető magyarázó változó kategóriái szerinti *felosztása* (splitting);
- az algoritmus addig folytatja rekurzív módon a kategóriák egyesítését és az esetek felosztását, míg el nem ér valamely előre definiált *megállítási* kritériumot (stopping).

A független változók kategóriáinak egyesítése

Az első lépésben a CHAID minden magyarázó változó esetében összevonja azokat a kategóriákat, melyek legkevésbé különböznek egymástól az m különféle kategóriával rendelkező célváltozó tekintetében. Ehhez X_i kategóriái közül az összes lehetséges módon kiválaszt kettőt. Amennyiben a vizsgált magyarázó változó K különböző kategóriával rendelkezik, a kiválasztás $K \times (K-1)/2$ féleképpen történhet. Ezt követően $K \times (K-1)/2$ különböző, $(2 \times m)$ méretű kontingenciátáblára Pearson-féle khi-négyzet teszt segítségével kiszámolja, hogy milyen p szignifikanciaszinten tekinthetők X_i kiválasztott kategóriapárjai és Y kategóriái függetlennek egymástól.

A következő lépésben kiválasztásra kerül az a kontingenciátábla, mely a *legmagasabb* p értékkel rendelkezik. Ezt az értéket az eljárás összeveti, a modellkészítő által előre lerögzített, $\alpha_{\text{egyesítés}}$ küszöbértékkel (a programcsomagok általában a szokásos 5 százalékos szignifikanciaszintet szokták felkínálni alapértelmezésként). Amennyiben $p > \alpha_{\text{egyesítés}}$ a kontingenciátáblázat X_i kategóriapárja egy új önálló kategóriába kerül egyesítésre. Ebben az esetben X_i eredeti kategóriáinak száma eggyel csökkent, és az algoritmus újból indul az elejétől, azaz az „új” kategóriapárok kiválasztásától (amelyek között nyilván lehetnek olyanok is, melyeket az előző ciklusban is kiválasztottak), az azokhoz rendelt kontingenciátáblákhoz tartozó p értékek kiszámolásáig.

A kategóriák összevonásának ciklusa mindaddig folytatódik, míg a legmagasabb p értékkel rendelkező kontingenciátáblára igaz nem lesz a $p > \alpha_{\text{egyesítés}}$ feltétel. Ekkor a vizsgált magyarázó változó (X_i) esetében a ciklus leáll, és az algoritmus a következő lépésben most már X teljes, lehetséges összevonások utáni, új kategória-struktúrájára számolja ki p értékét. A könnyebb eligazodás végett jelöljük az i magyarázó változó esetében ezt a szignifikanciaszintet $p_{x(i)}$ módon. Ezek után a modellkészítő igénye szerint kerül sor a $p_{x(i)}$ ún. Bonferroni-kiigazításra.¹ (Lásd bővebben a Függelékben.)

¹ A Bonferroni-kiigazítást több hipotézis egyidejű tesztelése során szokták alkalmazni. Amennyiben „ n ” féle különböző hipotézisünk van, melyeket külön-külön α szignifikanciaszinten tesztelnénk, belátható, hogy együttes fennállásuk esetén a szignifikanciaszintet α/n szinten kell megválasztani ahhoz, hogy az első fajú hiba elkövetésének valószínűsége ne legyen nagyobb, mint α . Esetünkben a különböző és egyidejűleg fennálló hipotéziseket a fastruktúra különböző szintjein vizsgált függetlenségi hipotézisek jelentik.

A CHAID minden magyarázó változó esetében végrehajtja a fent leírtakat, aminek eredményeképpen az összes X_i esetében megtörténnek a lehetséges kategóriaösszevonások, és minden magyarázó változó rendelkezik egy $p_{x(i)}$ (vagy kiigazított $p_{x(i)}$) értékkel.

A felosztás

A következő lépésben az X_i magyarázó változók közül kiválasztásra kerül a *legkisebb* „ $p_{x(i)}$ ” értékkel rendelkező. Ezt az értéket az eljárás összeveti, a modellkészítő által előre meghatározott, $\alpha_{\text{felosztás}}$ küszöbértékkel. Amennyiben $p_{x(i)} < \alpha_{\text{felosztás}}$, megtörténik az esetek felosztása X_i kategóriái szerint. A felosztás eredményeként a megfigyelések adatbázisa annyi részre esik szét, ahány (lehetséges összevonások utáni) kategóriával a felosztás alapjául szolgáló magyarázó változó rendelkezett. A felosztás utáni részadatbázisok fogják a fastruktúra következő szintjét jelenteni.² Ha $p_{x(i)} > \alpha_{\text{felosztás}}$, a felosztás nem történik meg, az adott szint tovább már nem bontható.

A megállás

Felosztás után az algoritmus az első pontnál (kategóriák egyesítése) újraindul, azzal a különbséggel, hogy most már az esetek felosztása után létrejövő részadatbázisokon külön-külön folytatódik a magyarázó változók kategóriáinak összevonása, majd az újbóli felosztás. A ciklusok (összevonás–felosztás) mindaddig tartanak, amíg el nem érik valamelyik megállási kritériumot. Ezek a következők lehetnek:

- $p_{x(i)} > \alpha_{\text{felosztás}}$;
- az esetek a magyarázó változók tekintetében nem különböznek egymástól (ugyanazon értékekkel rendelkeznek minden magyarázó változóra vonatkozóan);
- az esetek a célváltozó ugyanazon értékével rendelkeznek;
- a felosztandó részadatbázis esetszáma nem éri el a modellkészítő által előre definiált esetszámot;
- az újbóli felosztás során keletkező új részadatmátrixok valamelyikének esetszáma nem éri el a modellkészítő által előre definiált esetszámot;
- a felosztások száma eléri a modellkészítő által előre definiált számot (A fastruktúra szintjeinek száma = felosztások száma).

A leírtak szemléltetésére vegyünk egy példát, mely a *SPSS Answer Tree* számítógépes programcsomag segítségével készült. A példában egy hitelminősítési problémát találunk, melyben a rendelkezésre álló adatbázis segítségével szeretnénk kategorizálni a kérelmezőket aszerint, hogy mekkora hitelkockázatot jelentenek. Az adatbázis 323 esetet tartalmaz, négy, kategóriás kimenetelű magyarázó változóval. Ezek a következők:³

- X_1 : korosztály (fiatal, középkorú, idős) – Age Categorical (young, middle, old);
- X_2 : van-e AMEX kártyája (igen/nem) – AMEX card (yes/no);
- X_3 : fizetését hetente vagy havonta kapja (hetente/havonta) – Paid weekly/monthly (weekly pay/monthly salary);

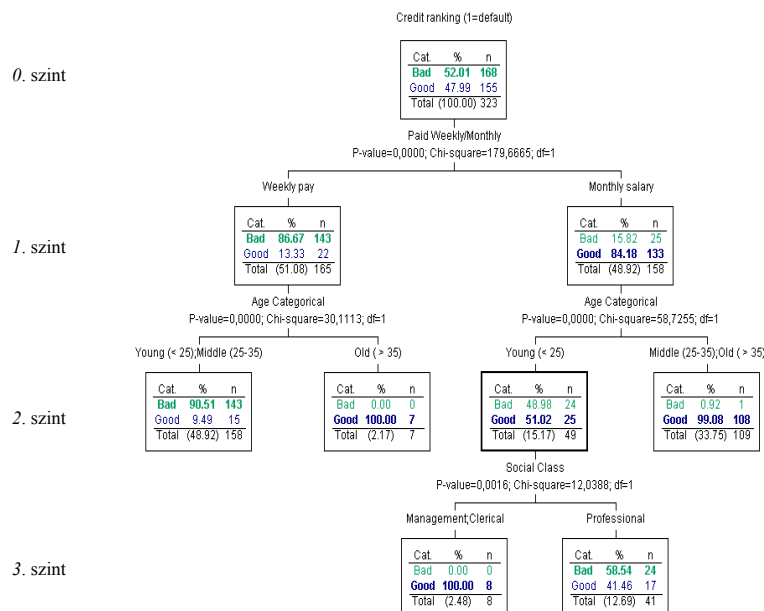
² Az első szint, maga a kiindulási adatbázis volt.

³ Az SPSS Answer Tree eredményei angol nyelven jelennek meg, ezért a felsorolásban magyarul és angolul egyaránt feltüntetjük az egyes változók neveit és kategóriáit.

– X_4 : foglalkozása (vezető, szabadfoglalkozású, irodai foglalkozású, szakmunkás, segédmunkás) – Social class (management, professional, clerical, skilled, unskilled).

Célváltozó (Y) értelemszerűen legyen a hitelbesorolás, melynek két kategóriája van: *jó* és *rossz*. A CHAID modell megépítésének első lépése, hogy lerögzítjük α egyesítés és α felosztás értékeit. Mindkettőnél a program által alapbeállításként ajánlott $\alpha = 0,05$ szintet fogadtuk el. A második lépés a megállási kritériumok meghatározása. A program a felosztások maximális számára vonatkozóan négyet javasol, ami a fastruktúrában legfeljebb öt szintet jelent. A felosztásra kerülő részadatbázisok minimális esetszáma 25, a felosztás során keletkező új részadatbázisok minimális esetszáma pedig 1.

Az alapparaméterek beállítása után az SPSS Answer Tree programcsomag a következő ábrán látható CHAID modellt alakította ki.



A 0. szinten látható a teljes adatbázis eloszlása a célváltozó kategóriái szerint. Egy-egy kis tábla jobb oldali oszlopa tartalmazza az egyes kategóriák elemszámait, a középső oszlop ugyanezek százalékos megoszlását, míg a bal oldali oszlopban láthatjuk feltüntetve az egyes kategóriákat. Ha csak ezt a változót ismernénk, és ennek alapján kellene döntünk egy hitelkérelemről, a legkisebb hibát akkor követnénk el, ha mindenkit elutasítanánk. Ekkor összességében a minta alapján az esetek 48 százalékánál hibát követnénk el. A CHAID segítségével a téves döntések aránya csökkenthető. Ehhez az algoritmus első részében minden egyes magyarázó változó esetében elvégzi a lehetséges összevonásokat, majd a magyarázó változók közül kiválasztja a legkisebb $p_{x(i)}$ értékkel rendelkezőt. Esetünkben ez a kérelmező fizetésének gyakoriságát leíró változó ($p_{x(i)} = 0,0000$, *khi-négyzet* = 179,6665). Ez azt is jelenti, hogy ennek a változónak a kategóriái gyakorolják a

legnagyobb befolyást a hitelkockázatra. A kiinduló adatbázis felosztása ennek a változó-
nak a kategóriái mentén történt meg. Ezzel elérkeztünk a fastruktúra 1. szintjére. A fel-
osztás eredményeképpen előálló két részadatbázisban újból megfigyelhetők a célváltozó
kategóriáinak eloszlásai az egyes részadatbázisokat reprezentáló táblácskákban. Látható,
hogy az ismérvkategóriához való tartozás ismerete a hitelezés kockázatával kapcsolatos
bizonytalanságunkat lényegesen csökkenti. Ha mindenkit, aki hetente kap fizetést, eluta-
sítanánk és mindenkinet, aki havonta kap jövedelmet hitelezni, akkor az esetek
 $(22+25)/(165+158) = 14,55$ százalékban döntenénk csak helytelenül az adatbázis által
reprezentált világban. A kezdeti 48 százalékos döntési bizonytalanságunkat $PRE = (48 -$
 $- 14,55) / 48 = 69,7$ százalékkal sikerült csökkenteni azért, hogy ismerjük a fizetések
gyakoriságát.⁴ A 14,55 százalékos bizonytalanság tovább csökkenthető, ha az összevonó-
felosztó algoritmust tovább folytatjuk és a fastruktúra 2. szintjére lépünk. Látható, hogy a
következő legnagyobb hatású (legkisebb $p_{x(i)}$ -szel rendelkező) magyarázó változó a kor-
osztály. A két részadatbázison a korosztály kategóriáinak összevonása másképpen történt
meg: a havi fizetéseknél a fiatalok, míg a heti fizetéseknél az idősek képeznek önálló ka-
tegóriát, míg a másik két korosztályi kategória összevonásra került. Döntési bizonytalansá-
gunk most már $(15+0+24+1)/(158+7+49+109) = 12,38$ százalékra csökkent
($PRE = 14,9\%$). A fastruktúrának ezen a szintjén már négy diszjunkt részadatbázisra lett
felosztva az eredeti adatmátrix. A fastruktúra heti fizetések ágán a korosztályi kategóriák
alapján felosztott két részadatbázist az algoritmus már tovább nem bontotta. Az „idős”
kategóriánál ez azért következett be, mert ennek a részadatbázisnak az elemszáma kisebb
a modell futtatása előtt beállított huszonöttnél, tehát az algoritmus itt elért egy megállási
kritériumot. A „középkorú–fiatal” összevont kategóriával jellemzett részadatbázis elem-
száma ugyan kellően nagy (158), de a $p_{x(i)}$ értékek egyike sem volt kisebb⁵ az előre beál-
lított $\alpha_{\text{felosztás}} = 5$ százalékos értéknél, így a felosztás nem következett be. A fastruktúra
havi fizetések ágán, a „fiatal” korosztály ágán tudott az algoritmus a fa 3. szintjére lépni.
A felosztás a foglalkozás szerint történt meg. Látható, hogy a vezető és az irodai dolgozó
kategória összevonásra került, és feltűnhet, hogy hiányzik a szakmunkás és a segédmun-
kás kategória. Ennek az a magyarázata, hogy az induló adatbázis nem tartalmazott olyan
esetet, melynél fiatal és havi fizetéssel rendelkező kérelmező szak- vagy segédmunkás
lett volna. Az alapparaméterek rögzített szintjén a fastruktúra kiépítése véget ért. A végső
struktúra segítségével a hitelkockázattal kapcsolatos döntési bizonytalanságunk 10,2 szá-
zalékra csökkent ($PRE = 17,6\%$). Az eljárást összefoglalóan mutatja a következő tábla.

A CHAID eljárás összefoglaló táblája

Szint	Hibás besorolások aránya (százalék)	PRE mutató (százalék)	
		az előző szinthez viszonyítva	a 0. szinthez viszonyítva
0.	48,00	–	–
1.	14,55	69,7	69,7
2.	12,38	14,9	74,2
3.	10,20	17,6	78,8

⁴ A PRE mutató a kapcsolat szorosságának általános mutatószáma, azt méri, hogy egy újabb változó bevonása a magyará-
zó változók közé hány százalékkal csökkenti a magyarázat bizonytalanságát. Ebben az esetben konkrét jelentése az, hogy a hi-
bás besorolások száma az egyes szinteken hány százalékkal csökken a korábbi szinten mérthez képest.

⁵ Itt már csak két magyarázó változó a „foglalkozás” és „Van AMEX kártyája” esetében kerül sor „ $p_{x(i)}$ ” érték számítására.

A tábla második oszlopa jól mutatja, hogy a növekvő szintek (növekvő számú magyarázó változó) hogyan eredményeznek egyre pontosabb besorolásokat. A harmadik oszlop a viszonylagos hibacsökkenést mutatja. Látható, hogy ez tendenciaszerűen csökkenő, de a csökkenés nem monoton. Végül az utolsó oszlop monoton növekvőn mutatja, hogy az induló állapothoz képest az egyes lépések után mekkora relatív hibacsökkenés érhető el. A végeredményül kapott 78,8 százalék jelentése az, hogy az összes szóba jöhető magyarázó változó együttesen közel 80 százalékkal csökkenti a hitelbesorolásban tapasztalt induló bizonytalanságot. Az ehhez tartozó döntési szabály tehát jó alapot nyújt a banknak a kérelmezők kockázat szerinti besorolására.

A beállítható paraméterek

Általában elmondható, hogy adott adatbázis esetén a fa összetettségét és mélységét (szintjeinek számát) alapvetően az határozza meg, hogy a futtatás előtt milyen értéken rögzítjük az alapparamétereket, melyek összefoglalva a következők:

- $\alpha_{\text{egyesítés}}$,
- $\alpha_{\text{felosztás}}$,
- felosztások maximális száma,
- felosztandó részadatbázis minimális esetszáma,
- a felosztással keletkező részadatbázisok minimális esetszáma.

Minél kisebb $\alpha_{\text{egyesítés}}$, annál több kategória egyesítésére számíthatunk a magyarázó változók esetében. Az $\alpha_{\text{felosztás}}$ kis értéke, viszont a felosztások számát, és ezáltal a fa összetettségét, csökkenti.

A CHAID TOVÁBBFEJLESZTÉSEI

A 80-as években a CHAID-et számos területen alkalmazták. A gyakorlat során merült fel az az igény, hogy a CHAID képes legyen mind a magyarázó változók, mind a célváltozó tekintetében mennyiségi ismérvek kezelésére is. A problémát a magyarázó változók esetében úgy oldották meg, hogy az algoritmus a mennyiségi változókat kategóriás változókká transzformálja oly módon, hogy X decilisei által meghatározott intervallumokat tekintik kategóriáknak. Amennyiben a célváltozó mennyiségi, az algoritmus khi-négyzet tesztek helyett F -teszteteket alkalmaz annak megállapítására, hogy milyen p szignifikanciaszinten tekinthetők a célváltozó X kategóriáirajai által meghatározott rész-átlagai azonosnak. Az alkalmazás során derült ki az algoritmus azon gyengesége is, miszerint a kategóriák összevonása során nem mindig éri el azt a kategória-struktúrát, melynél a p_x érték a legkisebb, azaz, amelynél a felosztás optimális (amennyiben az adott változó mentén történik az adatbázis felosztása). Ez annak a következménye, hogy az összevonási algoritmus leáll, amennyiben a megmaradt kategóriapárokat az algoritmus $\alpha_{\text{egyesítés}}$ függvényében statisztikailag függetlennek tekinti. A probléma orvoslására javasolta 1991-ben *D. Biggs*, *B. de Ville* és *E. Suen* az eredeti CHAID továbbfejlesztését, amelyet *Exhaustive CHAID*-nek neveztek el. Az *Exhaustive CHAID* csak annyiban kü

lőnbözik az eredeti CHAID-től, hogy az összevonási algoritmusnál nincs $\alpha_{\text{egyesítés}}$ összevonási és leállási kritérium. E helyett úgy dolgozik, hogy mindenféleképpen egyesíti az algoritmus első ciklusában azt a kategóriapárt, melynek a legmagasabb a p értéke, majd az így keletkezett új kategória-struktúrára kiszámolja a p_x -t. Az így nyert kategória-struktúrát a hozzá tartozó p_x -szel együtt „megjegyzi”. A következő lépésben az eljárás belép a második ciklusba és az előző kategória-struktúrán újból kialakítja a lehetséges kategóriapárokat és egyesíti azt a kategóriapárt, melynek legmagasabb a p értéke. Az új struktúrára megint kiszámolja a p_x értéket és a struktúrával együtt eltárolja a memóriájában. Az eljárás mindaddig folytatódik, míg csak két kategória marad. Ekkor az algoritmus visszamenőleg megkeresi azt a kategória-struktúrát, melyhez a legkisebb p_x érték tartozott. Ha majd az esetek felosztása ennek a változónak a mentén történik, a felosztás alapja az így létrehozott kategória-struktúra lesz. Ettől a ponttól kezdve minden ugyanúgy megy tovább, mint a CHAID esetében (ezután kerül sor a különböző magyarázó változókhoz tartozó p_x értékek összevetésére stb.).

Az Exhaustive CHAID hátránya a CHAID eljárással szemben az, hogy lényegesen számításgényesebb, aminek következtében, különösen nagy adatbázisok esetében a modell felépítésének időtartama jelentősen megnövekedhet. Ráadásul a gyakorlati tapasztalat azt mutatja, hogy sok esetben ugyanannak a problémának a megközelítésekor az CHAID és az Exhaustive CHAID ugyanazon fastruktúra kialakulásához vezet.

FÜGGELÉK

A Bonferroni-kiigazítás. Tegyük fel, hogy két hipotézisünk van, melyeket egyaránt α szignifikanciaszinten kívánunk tesztelni. Mindkét hipotézis esetén α jelenti az első fajú hiba elkövetésének valószínűségét, azaz a két hipotézis vonatkozásában ugyanazt az α szignifikanciaszintet határozzuk meg. Jelöljük A_1 -gyel az egyik és A_2 -vel a másik hipotézis esetében az első fajú hiba bekövetkezésének eseményét. Ekkor, a valószínűség-elmélet által használt jelölésekkel

$$P(A_1) = P(A_2) = \alpha.$$

Annak a valószínűsége, hogy legalább az egyik hipotézis esetében elkövetjük az első fajú hibát:

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2).$$

Jelöljük A_1 komplementerét A_{1c} -vel és A_2 komplementerét A_{2c} -vel. Ezek jelentik azt, hogy az első fajú hibák nem következnek be. Annak együttes valószínűsége, hogy sem az első, sem a második hipotézis esetében nem követjük el az első fajú hibát:

$$P(A_{1c} \cap A_{2c}) = 1 - P(A_1 \cup A_2) = 1 - P(A_1) - P(A_2) + P(A_1 \cap A_2).$$

Mivel tudjuk, hogy $P(A_1 \cap A_2) \geq 0$, felírható a következő, ún. Bonferroni-egyenlőtlenség két hipotézisre:

$$P(A_{1c} \cap A_{2c}) \geq 1 - P(A_1) - P(A_2) = 1 - 2\alpha.$$

Az általános formula n különböző hipotézis esetén:

$$P\left(\bigcap_{i=1}^n A_{ic}\right) \geq 1 - n\alpha.$$

Ha például tíz különböző hipotézissel dolgozunk egyszerre, melyek mindegyikét $\alpha = 0,05$ szignifikanciaszinten teszteljük, látható, hogy annak a valószínűsége, hogy egyik hipotézis esetében sem követjük el az első fajú hibát, nagyobb vagy egyenlő, mint 50 százalék. Ha az együttes valószínűsége a szokásos $\alpha = 0,05$ korlátot szeretnénk definiálni, akkor az egyedi (hipotézisenkénti) szignifikanciaszintet legfeljebb $\alpha/n = 0,05/10 = 0,005$ értéken szükséges rögzíteni.

IRODALOM

- BIGGS, D. – DE VILLE, B. – SUEN, E. (1991): A method of choosing multiway partitions for classification and decision trees. *Journal of Applied Statistics*, 18. sz. 49–62. old.
- GNANADESIKAN, R. (1977): *Methods for statistical data analysis of multivariate observations*. John Wiley & Sons. Inc. New York.
- KASS, G. (1980) An exploratory technique for investigating large quantities of categorical data. *Applied Statistics*, 29. évf. 2. sz. 119–127. old.
- MAGIDSON, J. – SPSS INC. (1993): *SPSS for Windows CHAID release 6.0*. SPSS Inc. Chicago.

SUMMARY

CHAID, or Chi-square Automatic Interaction Detection is a classification tree modelling technique. This exploratory data analysis method is used to study the relationships among a dependent variable and a large series of possible predictor variables and their interactions. The CHAID evaluates complex interactions of the predictors and the dependent variable, and displays the modelling results in an easy-to-interpret tree diagram.

A KANADAI STATISZTIKAI HIVATAL FEJLESZTÉSEI AZ EZREDFORDULÓN

A Kanadai Statisztikai Hivatal folyóiratának, a *Statistics Canada*-nak 2001. évi különszáma terjedelmes interjút közölt az ország főstatistikusával, dr. Fellegi Ivánnal, a közeljövő fő fejlesztési céljairól. A folyóirat sok éve tudatosan törekszik arra, hogy képet adjon a társadalmi–gazdasági folyamatok legfontosabb tényeiről és ezek összefüggéseiről. Az államháztartás különböző szintjei közötti jövedelemátcsoportosítás hatásvizsgálatát célozza például a tartományok szerinti („regionális”) gazdaságstatistika folyamatban levő fejlesztési projektje (Project to Improve Provincial Economic Statistics – PIPES).

Az interjúban a Hivatal vezetője részletesen kifejti a fontosabb statisztikai fejlesztések tartalmát. A kanadai törekvések közül azok érdemelnek figyelmet, amelyek hatást gyakorolhatnak a magyar felkészülésre is. Az interjú több témaköre érinti a magyar Központi Statisztikai Hivatallal (KSH) folytatott konzultációs, kooperációs együttműködést.

A bevezető rész bemutatja a statisztikusok felkészítésének és rendszeres szakmai képzésének kialakult formáit. A képzés az átalakuló adatgyűjtési, feldolgozási és elemzési feladatok gyorsan növekvő igényeihez igazodik. A Kanadai Statisztikai Hivatal elfogulatlan elemzéseinek fogadtatása kedvező, a mértékadó politikai körök elismerik és igénylik az adatokon alapuló megállapításokat például a termelékenység alakulásáról, az Észak-Amerikai Szabadkereskedelmi Egyezmény (NAFTA) külkereskedelmi hatásairól, az „agyelszívás” alakulásáról, a munka világáról, a jövedelmek egyenlőtlenségéről, az egészségi állapotról.

Fellegi Iván kiemeli a statisztikai hivatal és a kutatás, valamint az egyetemi oktatás közötti együttműködés jelentőségét. Az ország társadalom- és humántudományi kutatási tanácsa (Social Sciences and Humanities Research Council – SSHRC) ajánlásainak megfelelően a nagyobb egyetemeken kihelyezett „kutatói adatfeldolgozó központokat” (Research Data Centres) hoztak létre. A statisztikai

hivatal „fiókintézeteiként” működtetik ezeket a kutatói munkaállomásokat, amelyeken (a statisztikai törvény által előírt bizalmas kezeléssel) az intranetnek megfelelő jogosultsággal, bármikor helyben elérhetők az összes tárolt statisztikai információk.

A kanadai kormány Government On-Line (GOL) elnevezésű internetes fejlesztési programja révén, 2004 végéig, a világ tetszőleges pontján elérhető kormányzati informatikai rendszert alakítanak ki. Ehhez a statisztikai adatsorokat a közelmúltban indított CANSIM II internetes adattárház szolgáltatással csatlakoztatják. Az interjú részletesen kifejti a GOL és a statisztikai tájékoztatás fejlesztéseinek szerves összefüggéseit.

A CANSIM II 2001-től elérhető, és a legfontosabb összefoglaló statisztikai adatok ingyenesen megtekinthetők, azonban a részletesebb adatok csak térítés ellenében érhetők el. A hivatal a CD-lemezekon kiadott adatait a CANSIM II táblázatainak megfelelően állítja össze. A CANSIM II mind a www.statcan.ca saját honlapról, mind a kanadai kormány portáljáról elérhető, az intranet útján pedig minden statisztikus kétirányú kapcsolatban van az adattárházzal, jogosultságának megfelelő írással és olvasással. Az új szolgáltatás az előző CANSIM-verzió koncepciójától abban tér el, hogy a CANSIM II nemcsak a statisztikai adatsorok tárolására alkalmas, hanem több szempontú kapcsolatok (cross-classified data) érvényesítésére is.

A rendszer konkrét igények alapján magából a statisztikai adatbázisból állítja össze a kért (változtatható részletességű) adatsorokat. Ennek révén lényegesen csökkenthető az adatközlés hibája és időigénye. A kutatók igényére tömörítvényeket készítettek, és kulcsszavak szerint is kereshetők az interneten elérhető adatsorok. Ezek az információk (Research Papers) mindkét kanadai hivatalos nyelven (angolul és franciául) elérhetők. A metaadatok integrált adatbázisát (Integrated Metadata Database – IMDB) 2000. novemberben helyezték üzembe. A

látogató ebből a modulból kattintással léphet át a a megfelelő katalógusba (Online Catalogue) vagy az éppen keresett adatsorhoz.

A hivatal üzleti tervei szerint évente 1 millió kanadai dollár bevételt hoz a CANSIM II szolgáltatás. Egy adatsor ára 3 dollár a továbbiakban is, ezt a CANSIM felhasználói már széles körben elfogadták. Az árképzés elfogadható mértékeit tesztelték, mintegy 100 felhasználó bevonásával. Olyan kormányzati előírás is elképzelhető, hogy tegyék ingyenessé ezt a részletezett statisztikai tájékoztatást, amit a hivatal úgy teljesíthet, ha ennek megfelelő térítéshez jut a költségvetésből, a GOL alapján.

A hivatal honlapjával 2000 októberében naponta 15 000 kapcsolatot létesítettek, ez az 1999. októberi napi forgalom másfélszerese. A tanulók aránya az összes kapcsolatban 20 százalék volt 1999-ben, és 2000-re arányuk már közel 40 százalékra nőtt. A látogatók kifejtetik a véleményüket, elégedettségüket.

Az interjú válaszai összefoglalóan ismertetik a legfontosabb 2001. évi statisztikai adatgyűjtéseket. A 2001. május 15-i állapotnak megfelelő népszámlálás (2001 Census) kérdőívét a kormányzati szervek egyetértésével állították össze, számos új kérdéssel. A házastársi és élettársi kapcsolatok körében az azonos neműek viszonyát is felméri. Új kérdés a munkahelyen használt nyelv, a szülők születési helye. Ez a kérdéskör a bevándorlók második nemzedéke esetén ad képet a társadalmi-gazdasági helyzetükről.

A demográfiai felvételek közé tartozik az őshonos lakosság összeírása (2001 Aboriginal Peoples Survey), valamint az egészségükben, tevékenységükben korlátozottak adatfelvétele (2001 Health and Activity Limitation Survey). A szaktárcával egyeztetve a népszámlálást követően, 2002 márciusában kerül sor az etnikai csoportok adatfelvételére (Ethnic Diversity Survey). Az adatok alapján kellő támogatást szerezhetnek a 2006. évi lakossági összeírás (2006 Census) tartalmának kialakításához. A korábbi években megkezdett program folytatásaként tervezik a bevándorlók longitudinális felvételét (Longitudinal Survey of Immigrants).

A hivatal programjai között szerepel az oktatásügyi statisztikai mutatórendszerének fejlesztése (Pan-Canadian Education Indicators Program), figyelembe véve a szövetségi, a tartományi és a helyi tervezési, elemzési és folyamatos közszolgáltatási feladatok sajátosságait. A módszertani fejlesztések eredményei 2001-ben jelennek meg. Az Education Quarterly Review 2000. májusban adta közre az „agvelszívás” és „megnyert agyak” alakulásáról készített tanulmányt (Brain drain and brain gain). Az interjú bemutatja a kilencvenes évek adatsorainak fontosabb megállapításait.

Bekapcsolódtak a „diákok vizsgálata” (az OECD által koordinált) programba (Programme for International Student Assessment – PISA). Ez a program több mint 30 országra kiterjedően, három nagy szakaszban irányoz elő felmérést (2000-ben, 2003-ban és 2006-ban, különböző kérdéskörökkel).

Kanadában, az említett OECD-programhoz kapcsolódva, 2000-ben indult a pályakezdő fiatalok felvétele (Youth in Transition Survey – YITS), két nagy korcsoportra: az egyikbe a 15 évesek, a másikba a 18-20 évesek tartoznak. Az említett YITS-felvétel két évenként ismétlődik, amíg ezek a pályakezdők betöltik a 25-29 életévüket. Új, tízezer fős mintával kezdik meg a gyermekek és fiatalok longitudinális felvételét. (National Longitudinal Survey of Children and Youth).

A korábbi években a háztartások pénzügyi helyzetét a Survey of Consumer Finances (SCF) felvétellel figyelték meg, ezt felváltotta az 1999. júniusban indított, a munka és a jövedelem alakulását megfigyelő adatgyűjtés (Survey of Labour and Income Dynamics – SLID). Az interjú kifejti a jövedelemfelvételek kiadványainak fontosabb megállapításait, amelyek eddig 13 különböző sorozatban jelentek meg. Az új adatgyűjtésre alapozva egyetlen összefoglaló kiadványt (Income in Canada) indítottak.

A tervek szerint 2002-ben jelennek meg a háztartások felszereltségéről, tartozásairól szóló felvétel eredményei (Household Survey on Assets and Debts), amely alapján kép alkotható a kanadai háztartásokban felhalmozott vagyonról. A negyvenes évek második felében és az ötvenes évek elején született nemzedék (ez az ún. „baby boomers” korosztály, amely közeledik a nyugdíjba vonuláshoz) tetemes nyugdíjvárománnyal rendelkezik, és ez az „eszköz” sok kanadai számára különösen jelentős.

A Hivatal az egészségügyi statisztika eredményeit a „Health Care in Canada 2000: A first annual report” című kiadványban adta ki, amelynek megállapításait a cikk részletesen bemutatja. Longitudinális népegészségügyi felvételt végeznek (National Population Health Survey – NPHS), ennek 4. ciklusa 2000. júniustól 2001. márciusig tartott. A cikk részletezi a longitudinális népegészségügyi felvételek eredményeit. Az egyes települések egészségügyi ellátottságát vizsgáló felvétel (Canadian Community Health Survey – CCHS) első ciklusa 2000. szeptemberben indult és 2001. szeptemberéig tart. Viszonylag nagy, 130 000 elemű mintát választottak, összesen 136 kanadai egészségügyi körzetben. A kombinált adatgyűjtési ismérvek alapján sokoldalúan hasznosítható adatok publikálhatók, az első elemzések 2001 tavaszán jelennek meg.

A tartományok szerinti („regionális”) gazdaságstatisztika fejlesztési projektje (PIPES) előkészítése több millió dollárt igényel, és több évig tart. Már 1999-ben elkészült az ország 13 területi egysége szerint tagolt 1996. évi input-output tábla. A projekt részeként hajtották végre 2000-ben az 1997. évi bázisra való áttérést. A tartományi részletezésű adatgyűjtések fontos kiegészítéseket adnak a PIPES koncepciójának megvalósításához.

A háztartások kiadásaira (Household Spending Survey), az utazásokra (Canadian Travel Survey), a javításra és felújításra (Repair and Renovation Survey) vonatkozó adatfelvételek korszerűsítése is része a PIPES-koncepciónak. Kísérleti felvétel indult az egységes vállalati adatgyűjtésről (Unified Enterprise Survey – UES), ennek első eredményei 2000 elején jelentek meg, olyan kiválasztott tevékenységekre, mint például a személyszállítás (taxizás), az üzenet- és futárszolgálat, ingatlanok bérbéradása és közvetítése, étel- és italszolgáltató helyek üzemeltetése, strandszolgáltatás és hasonlók.

Sok új ismeretet nyertek a kísérleti eredmények feldolgozása során, amelyeket a tartományi gazdasági számlák (1997. tárgyévre vonatkozó) összeállítása során hasznosítanak. Folytatják ezt az egységes (UES) felvételt a nagy- és kiskereskedelemről, a vezetői tanácsadásról, a fordításról, a munkaügyi szolgáltatásról, a lapok adatbázisairól és szakosodott kiadásairól is.

A PIPES-fejlesztés révén kiterjesztik, aktualizálják a statisztikai regisztert, amely jelenleg közel kétmillió gazdálkodó egységet tartalmaz. Amikor megkezdődtek a PIPES előkészületei a statisztikai regiszter nem egészen egymillió szervezetet tartalmazott. A regiszterben az újonnan bevezetett Észak-Amerikai Ágazati Osztályozási Rendszer (North American Industry Classification System – NAICS) szerinti tevékenységi besorolásokat is alkalmazzák.

Korábban a Standard Industry Classification (SIC80) ágazati osztályozáson alapultak a gazdaságstatisztikai közlések, legutóbb 1992. évi bázisúval. Az észak-amerikai NAICS ágazati osztályozásra 1998-tól tértek át, egyidejűleg az új bázisú 1997 lett.

A regiszter minőségének javítása kettős feladattal jár. Egyrészt azonosítani kell a nem működő egységeket, másrészt a tényleges tevékenységeknek megfelelően kell pontosítani a szervezetek ágazati besorolását. Ezzel a regiszter jobban megalapozza a mintavételeket. Nagy erővel fejlesztik a „felhasználóbarát” adatbázis-kezelő programokat, hogy az egyes statisztikai feladatok gazdái közvetlenül elemezhesék a mintaválasztás aktuális keretsokaságát, elkészíthessék azokból a céljaiknak megfelelő megoldásokat.

Korszerűsítik a kanadai környezeti statisztikát, országos környezeti információs rendszert alakítanak ki, és szakértői csoport (Task Force on the Canadian Information System on the Environment) dolgozik a fejlesztés vezetői döntéseinek előkészítésén. A cél az, hogy kialakítsák a fenntartható fejlődés megfelelő statisztikai mutatóit. Például a Kiotóban, 1997. decemberében elfogadott dokumentum értelmében Kanada 2010-ig 6 százalékkal csökkenti az üvegházhatást okozó gázok emisszióját az 1990. évihez képest. A statisztikusok feladata az ehhez szükséges beszámolási rendszer kialakítása az energiát termelők és felhasználók megfigyelésére.

Az Internet használatának 1999. tárgyévre vonatkozó éves felvételeinek (Household Internet Use Survey) eredményei 2000. májusban jelentek meg. Az interjú kifejti ennek a kiadványnak a fontosabb megállapításait. Figyelemmel kísérik a feldolgozóipar innovációit, amelynek felvételeit a hivatal az ipari, a környezeti és a tudományos kutatásokért felelős kormányzervekkel együttműködve szervezi.

Különös figyelmet kapnak az információs és távközlési technológiák átlagosnál gyorsabb fejlesztései. Kanada szempontjából nagy jelentőségű az erdőgazdálkodás innovációja, amelyre a szaktárcával együttműködve most készítenek elő közös adatgyűjtést (a partner a Natural Resources Canada).

A hivatal, az ipari szaktárcával együttműködve 2000 nyarán készítette elő a mikro- és kisvállalatok adatgyűjtését, a legfeljebb négy főt foglalkoztató gazdálkodó szervezetek körében. Külön programot indítottak a közepes és kisvállalatok pénzügyi helyzetének felvételére, ezen belül a finanszírozás lehetőségeire, jellegzetes forrásaira. Egyrészt közvetlenül megfigyelik a kisvállalatok pénzügyi keresletét, másrészt a pénzügyi intézetek oldalán vizsgálják a kínálat alakulását. A próbák értékelését követően 2001 közepén postázzák az adatgyűjtés kérdőíveit.

Három észak-amerikai ország együttműködésével adták ki a földrész szállítási statisztikáit összehangoltan bemutató kiadványt, „North American Transportation in Figures” címmel. A közlekedési szaktárcával (Transport Canada) együttműködve olyan új adatgyűjtést (Canadian Vehicle Survey) készítenek elő, amely a közúti szállítás adatainak meglévő hiányait rövidesen pótolja. A tartományokban bejegyzett közúti járművek igazgatási nyilvántartásai szolgáltatják a minta kiválasztásának keretsokaságát, és arra kéri az adatszolgáltatókat, hogy vezessenek a közúti szállítás adatait tartalmazó „útnaplót”. Ez a módszer a világon egyedülálló. A szaktárcától veszik át az utak és a járművek fizikai állapotára, a biztonságos közúti közlekedés feltételeire és az üzemanyag-fogyasztásra, valamint a

környezet terhelésére vonatkozó közlekedéssziszti-
kai alapadatokat.

A hivatal 2000. márciusban adta ki először azoknak a szervezeteknek az adatgyűjteményét, amelyek az 1993 és 1997 közötti időszakban exportot teljesítettek. A sorozat az 1998. és 1999. évi exportban részt vevők adataival folytatódik. Az exportálók adatai kiegészítik a külkereskedelmi forgalom árucsoportok szerinti havi kiadványait. Az interjú ismerteti a megjelent adatgyűjtemény fontosabb megállapításait.

A kanadai adó- és vámszervek (Canada Customs and Revenue Agency – CCRA) most alakítják át a külkereskedelmi adatok megfigyelését; változnak mind a behozatal, mind a kivitel okmányai, amelyek eddig megalapozták a nemzetközi áruforgalom statisztikáit. A statisztikáknak alkalmazkodniuk kell azokhoz a vámkezelési eljárásokhoz, amelyeket a jogszabályok előírnak, azonban olyan tárcaközi egyeztetéseket folytatnak, amelyek révén érvényesíthetők az új okmányok statisztikai feldolgozásának szempontjai.

Kiemelt fejlesztési feladat az adóhatóság általános forgalmiadó-állományainak statisztikai célú hasznosítása. Ehhez rendszeres és megbízható adatátadást kell előkészíteni, a fogadó statisztikai adatbázisok megfelelő kialakításával. Elsősorban a szolgáltatási statisztikák minőségét lehet javítani ezzel a kiegészítő információval, és csökkenthetők az adatszolgáltatási terhek. A kereskedelem statisztikáinak fejlesztéséhez fontos támpontot szolgáltatnak az adóhatóságtól átvett adatok.

Az interjú nagy figyelmet fordít a statisztikai elemzésekre, nemzetközi összehasonlításokra. Az Egyesült Államok és Kanada statisztikai adatai alapján például elemezték a jövedelmek szintjét, valamint kialakult egyenlőtlenségeit, rámutatva a lényegesebb társadalmi összefüggésekre. A „tudásra alapozott gazdaság”, valamint az innovációk tényeit több kiadvány elemzi. Új monográfia elemzi a termelékenység alakulását Kanadában, az Egyesült Államok megfelelő mutatóival összehasonlítva.

Becsléseket készítenek a szoftverek beruházásaira, mind a kereskedelmi forgalmazású („vásárolt”), mind a saját fejlesztésű számítástechnikai programok körében. Az eddig folyó termelési felhasználásként elszámolt beruházás adataihoz az Egyesült Államok és más fejlett országok megfelelő koncepcióit is figyelembe veszik. A bruttó hazai termék (GDP) felhalmozási adatait ennek megfelelően pontosítani szükséges, a történelmi idősorokra alapozva.

A nemzeti számlákhoz szatellit számlaként kapcsolják az „önkéntes szektor” szervezeteinek számláit, azok bevételeit, kiadásait, felhalmozott eszkö-

zeinek nettó értékét, pénzügyi számláit, követeléseit és kötelezettségeit. Ilyen szervezetek esetén díjazás nélküli önkéntesek is végeznek munkát, ezért a „nem piaci tevékenységek” értékelését is meg kell oldani. Ilyen becslésekhez a nemzetközi szabványokat veszik alapul, hogy a nonprofit szervezetek teljesítményei is összehasonlíthatók legyenek a hagyományos nemzeti számlákkal (a GDP termelésével), valamint a nettó érték és a beruházás adataival.

Nagy feladat lesz a következő években az államháztartási számlák átalakítása, mivel eddig a pénzforgalmi adatokat vették alapul, ezután pedig a ráfordításokat tekintik az elszámolások alapjának. Más szemlélettel készülnek ezt követően például a tőkefelhalmozások államháztartási, valamint a szektorközi kiegyenlítő fizetések elszámolásai.

A statisztikai munka szervezeti hátterében jelentős változások kezdődtek, amelyeket az interjú részletesen kifejt. Fejlesztéseket hajtottak végre az informatikai környezet korszerűsítése érdekében.

A jövőben nagy erőket összpontosítanak az elektronikus adatgyűjtés feltételeinek megeremtésére, különös tekintettel a vállalati adatszolgáltatók növekvő érdekltségére és ilyen jellegű válaszolási készségére. Amennyiben a GOL-konceptió megvalósul, a statisztikai szolgálatnak is fel kell ajánlania ilyen adatszolgáltatási lehetőséget, amely a korábbiaknál lényegesen hatékonyabb, legalábbis a megfelelő számítástechnikai felkészültséggel rendelkező partnerek körében.

A Kanadai Statisztikai Hivatal technikai segítséget nyújt Kínának, Magyarországnak, valamint a bagladeshi, a dél-afrikai, az eritreai, az indonéz, a jamaikai, a kubai, a malaysiai, a fülöp-szigetekbeli, az orosz, az örmény, a török és a zambiai statisztikai hivataloknak. Megemlíti az interjú a Fellegi–Ryten páros elemző tevékenységét is, amelyet a svájci és a magyar statisztikai szolgálatok tevékenységének értékelése kapcsán végeztek.

A Kanadai Statisztikai Hivatal 1999-ben vezette be (négy önkéntesen vállalkozó szervezeti egység bemutatkozása alapján) a programok kétévenként esedékes beszámolóit (Biennial Program Report – BPR). A bemutatók rendkívül tanulságosak voltak mind a hivatal távlati terveit elbírálók, mind az összefoglaló jelentést készítők számára. Ezek kapcsán felülvizsgálták az egykor elfogadott program tervezett és megvalósult beruházásait, lehetőség nyílt az aktuális javaslatok kifejtésére. A tervek szerint négyévenként intézményesítve sor kerül az összes fejlesztési program hasonló beszámolójára (Quadrennial Program Report – QPR).

Nagy nyomás nehezedik a Kanadai Statisztikai Hivatalra, hogy a rövid távú gazdasági mutatók az

előírt időpontokban megjelenjenek, és erősödjenek a konjunktúra alakulására utaló jelzésük. Erős érdekek fűződnek ezzel párhuzamosan ahhoz, hogy csökkentsék az adatszolgáltatók terhelését. Minden szervezeti egység áttekintette lehetőségeit e feladatok megvalósítására, és sok kiváló ötletet vetettek fel. Ilyen javaslat például az önkéntes adatgyűjtések al-

kalmazásának szélesítése, a kérdőívek rövidítése, tartalmuk állandósítása. A legkisebb gazdálkodó szervezeteket ki kell hagyni a minta keretsokaságából, és az elérhető igazgatási adatállományok alapján kell becsléseket végezni, külön statisztikai adatfelvételek nélkül.

Nádudvari Zoltán

MAGYAR SZAKIRODALOM

HUNYADI LÁSZLÓ:

STATISZTIKAI KÖVETKEZTETÉSELMÉLET
KÖZGAZDÁSZOKNAK

Központi Statisztikai Hivatal. Budapest, 2001. 483 old.

A Statisztikai módszerek a társadalmi és gazdasági elemzésekben című sorozatban elsőként megjelent kötet – mint címe is jelzi – elsősorban közgazdászok számára készült, de haszonnal forgathatja minden olyan társadalomtudományi kutatással foglalkozó szakember, aki szeretne a statisztikai következtetésemélet kérdéseivel közelebbről megismerkedni. A könyv mindenképpen hiánypótló, hiszen a magyar szakirodalomban eddig nem találkozhattunk a témakört ilyen alaposan tárgyaló művel, amelyben a statisztikai következtetésemélet szerteágazó területéről elsődlegesen a közgazdasági, társadalomtudományi kutatások során felmerülő problémáknál hasznosítható elméleti kérdések és alkalmazások kerülnek bemutatásra.

A könyv nem alkalmazói kézikönyv. A szerző nem egyszerű receptkönyvet ad, amit csak ki kell nyitni a megfelelő helyen, és máris megtaláljuk az adott kérdésre a kész választ. A bölcs indián sem halat ad a barátjának, hanem inkább megtanítja, hogyan fogjon magának, ha éhes. Itt is inkább arról van szó, hogy ahol csak lehet, megmutatja a szerző a különböző módszerek, eszközök eredetét, működését. Részleteiben is ismerteti a különböző tesztek döntési szabályait vagy legalább levezetéseinek főbb lépéseit. Ez a módszer lehetővé teszi, hogy ha egy szokatlan problémával találkozunk, akkor is alkotó módon tudjunk hozzáállni, hiszen, ha valaki ismeri az elmélet alapelveit és azok gyakorlati alkalmazásának lényegesebb példáit, szükség esetén képes saját problémájára megfelelő, testre szabott választ találni.

A könyv tárgyalásmódja alapján valahol a hagyományos verbális statisztika és a matematikai statisztika határán áll. A szerző a hangsúlyt a statisztikai elvekre és alkalmazásokra helyezi, de megmutatja ezek matematikai hátterét is, ami bizonyos

esetekben segítheti a különböző összefüggések átlátását, a lényeg megértését. Mindezt úgy teszi, hogy a matematikában kevésbé jártas olvasót sem riasztja el az olvasástól, hanem mindenütt ajánl megfelelő szakirodalmat a vonatkozó ismeretek felfrissítésére vagy éppen elmélyítésére.

A könyv didaktikus felépítésének köszönhetően kiválóan alkalmazható tankönyvként, akár kurzusokhoz kapcsolódva, akár önálló tanuláshoz is. Valamennyi módszer tárgyalásakor elméleti és gyakorlati példák segítik az elmélet megértését és elsajátítását. Az esettanulmányok pedig biztosítják, hogy az itt szerzett ismeretek nemcsak elméleti tudásunkat gyarapítsák, hanem a gyakorlatba is átültethető, alkalmazható hasznos módszereket adjanak. Minden fejezethez kapcsolódnak gyakorló feladatok, amelyek az alaposabb megértést és az önálló problémamegoldást hivatottak segíteni (félreértés ne essék: nem feltétlen egyperces ujjgyakorlatokról van szó).

A szerző a következtetésemélet azon területeit tárgyalja, amelyek lényeges ökonometriai és közgazdasági alkalmazásokkal rendelkeznek, vagy kellemes, hogy rendelkezzenek.

A könyv a következtetésemélet módszertani alapjainak ismertetésével indul. A szerző az I. fejezetben az algebrai, valószínűség-számítási és statisztikai alapismeretekre támaszkodva összefoglalja mindhárom területről azokat a fogalmakat, eszközöket, módszereket, amelyekre a következtetésemélet tárgyalása során szükség lesz.

A lineáris algebrai összefoglaló sorra veszi azokat a jellegzetes mátrixokat, amelyeket a statisztikai elemzések során a leggyakrabban alkalmazunk, majd röviden áttekinti az ezekkel kapcsolatos leggyakoribb műveleteket, feladatokat. Ilyen a mátrixok spektrálfelbontása, deriválása vagy a mátrixok nyoma és determinánsa.

A valószínűség-elméleti áttekintés foglalkozik a valószínűségi változók momentumaival és azok tulajdonságaival, a fontosabb egy- és többdimenziós eloszlásokkal, valamint kiterjed a valószínűség-számítás aszimptotikus elméletére is. Ez utóbbira elsősorban nagymintás közelítések esetén lehet szük

ségünk, hiszen a mintaelemek függvényeinek tulajdonságairól gyakran csak határértékben, azaz végtelenbe tartó mintanagyság esetén lehet érdemi megállapításokat tenni. Ezzel kapcsolatban találkozhatunk több kevésbé ismert elméleti összefüggéssel és fogalommal, úgymint a sztochasztikus és a négyzetes konvergencia, a valószínűségi határ, az eloszlásbeli konvergencia, a határeloszlás és az aszimptotikus eloszlás.

A statisztikai bevezető inkább látókörbővítő jellegű. Rövid áttekintést ad a statisztika fontosabb területeiről, valamint összefoglalja a statisztika lényeges alapfogalmait, eredményeit és az alkalmazott jelöléseket. Itt ismerkedhetünk meg a likelihood függvény fogalmával, amely a statisztikai következtetésemélet egyik legfontosabb eleme, hiszen a klasszikus becslésemélet és a hipotézisvizsgálat is erre épül.

A módszertani alapok áttekintése után a II. részben rátér a szerző a pontbecslések témakörére. A becsléseméleti alapfogalmak tisztázását a pontbecslések kismintás és nagymintás kritériumainak bemutatása követi. Miután megvizsgáltuk, milyen a jó becslőfüggvény, azt is megtanulhatjuk, hogyan készíthetünk ilyeneket. A pontbecslés módszerei közül négyet mutat be részletesen is a könyv, ezek: a legkisebb négyzetek módszere, a khi-négyzet módszer, a momentumok módszere és természetesen a maximum likelihood módszer. Külön fejezetet szentelt a szerző az ökonometriában használt pontbecslések és azok tulajdonságainak bemutatására. A korábban megismert elvek és módszerek alkalmazásait pedig különböző példákkal szemlélteti.

Szintén egy fejezet szól a rendezett minták elméletéről, amelyről alapozó statisztikai könyvekben nemigen olvashattunk. Bemutatja a rendezést mint statisztikai műveletet ismert sokaságok elemeire, majd a minták rendezésével és a mintaelemek függvényeivel foglalkozik. Ezután a rendezett mintaelemek kismintás eloszlási eredményeit és a lényegesebb aszimptotikus eredményeket mutatja meg. Vizsgálja a rendezett mintákon alapuló becslések tulajdonságait, és bemutatja ezek alkalmazási lehetőségeit. Nem konkrét alkalmazásokról van itt szó, hanem lehetőségekről, ötletekről, hiszen a rendezett mintás statisztikák alkalmazása a közgazdasági gyakorlatban egyáltalán nem terjedt el.

A valóságban a társadalmi, gazdasági jelenségek vizsgálatánál gyakran nem tekinthetjük végtelen nagyok alapsokaságunkat. Ha a sokaság közelítőleg sem tekinthető végtelennek, ha nem ismerjük a sokasági eloszlást, ha a mintaelemek nem tekinthetők független azonos eloszlásúnak, mind-mind módosítják a statisztikai következtetésemélet egyes lépéseit, illetve

eredményeit. Szükség van ezért a véges sokaságokból vett minták és becslések sajátosságainak vizsgálatára is. A fejezet foglalkozik a mintavétel néhány kérdésével, és az esettanulmány a kétfázisú mintavétel egy alkalmazási lehetőségét mutatja be.

Előfordulhat, hogy nem ismerjük a sokasági eloszlást, ráadásul a megfigyelési adatok is tartalmazhatnak hibákat. Ezek a problémák vetették fel az igényt a robusztus és nemparaméteres eljárások iránt. A témakörhöz kapcsolódó fejezet az alapfogalmakon kívül a robusztusság mérésével és a robusztus becslőfüggvények készítésére szolgáló módszerekkel foglalkozik.

A pontbecslések áttekintése után, a III. részben az intervallumbecslések és a hipotézisvizsgálat témakörével foglalkozik a könyv. Az intervallumbecslés fogalmának és tulajdonságainak tisztázása után külön tárgyalja a kis- és a nagymintás konfidencia-intervallumok készítésének sajátosságait. A valóságban gyakran több paraméter együttes becslése és konfidencia-intervallumának meghatározása a feladat, aminek végrehajtása független változók esetén nem különösebben bonyolult, nem független változók esetén viszont korántsem egyszerű. Mindkét esetre találnunk megoldási módszereket, példákkal illusztrálva. Intervallumbecslések esetén a robusztusságot az eloszlásmentesség jelenti. Eloszlásmentes intervallumok készítésére alkalmas módszereket is bemutat a könyv. A konfidencia-intervallumok tárgyalása után a tolerancia-intervallumok fogalmával, fajtáival is megismerkedhetünk.

A becslésemélet áttekintése után a következtetésemélet másik nagy területére, a hipotézisvizsgálatra kerül sor. Ez a fejezet főleg a statisztikai próbák származtatására, készítésére koncentrálna. Olyan általánosan használható elveket és módszereket mutat be, amelyek segítségével jó tulajdonságú próbák készíthetők. Az OC görbe és az erőfüggvény megismerése után a jó próbák kritériumait (torzítatlanság, erő, konzisztencia) tekinthetjük át. A fejezet ismerteti a legerősebb próbát eredményező Neyman–Pearson-próbakészítési elvet, ebből levezeti az aránypróba egy egyszerű esetét, valamint az össze-tett nullhipotézisekre kiterjesztett általános likelihood próbát. Nagy minták esetén eloszlás-függetlenül konzisztens próbákat eredményeznek az aszimptotikus próbák. Három ilyen általánosan használható tesztet mutat be a könyv: a likelihood arány tesztet, a Lagrange-multiplikátor tesztet és a Wald-próbát.

Ismét külön fejezet szól az ökonometriában alkalmazott általános és speciális tesztekéről. Találunk példákat a normalitás és a regressziós modellek tesztelésére, a heteroszkedaszticitás és az autokorreláció tesztelésére, a sztochasztikus idősor-

modellek tesztjeire, a modellspecifikáció és a modellválasztás tesztjeire.

A könyv IV. része kitekintést nyújt a hagyományostól eltérő szemléletű bayesi statisztika alapjai felé. A bayesi elméletben a becslés tárgya nem rögzített mennyiség, hanem valószínűségi változó. A bayesi statisztika a következtetések kialakításánál a mintán kívül külső információkat is kiterjedten felhasznál. Nem feltételezi az ismételt mintavétel lehetőségét, ugyanakkor lehetővé teszi a szubjektív megítélések, vélemények valószínűségi eszközökkel történő beépítését a modellbe.

Napjaink információorientált világában egyre nagyobb szükség van olyan módszerekre, amelyek képesek többféle információ egyidejű, komplex hasznosítására. A számítástechnikai eszközök fejlődése szintén hozzájárult a számításigényes bayesi statisztika gyakorlati alkalmazásainak elterjedéséhez. A könyv áttekinti az alapfogalmakat (prior, posterior, Bayes-tétel...) és egyszerű példákkal bemutatja a bayesi következtetéselmélet főbb lépéseit, sajátosságait. A

bayesi módszerek alkalmazhatók szinte minden hagyományos statisztikai probléma kezelésére, ökonometriai becslésekben, és a statisztikai döntéseméletben is fontos szerepet játszanak. Alkalmazásukra a szerző gyakorlati példákat mutat be.

A könyv, amelyet *Mihályffy László* és *Rappai Gábor* lektorált, jól kiegészíti az egyetemen már megszerzett statisztikai alapismereteket. Ajánlható minden olyan szakember számára, aki a témakörbeli tudását bővíteni, elmélyíteni szeretné és a gyakorlatban is jól alkalmazható módszerek birtokába kíván jutni. A könyv olvasása után úgy ismerhetjük meg a következtetéselmélet rejtelseit, hogy ha egy addig ismeretlen feladat kerül elénk, akkor is önállóan gondolkodva, alkotó módon tudunk hozzáállni. Erre szükség is van, mert a valóságban korántsem egyszerű problémákkal találjuk magunkat szemben.

Oravecz Beatrix

STATISZTIKAI HÍRADÓ

SZEMÉLYI HÍREK

Elnöki dicséret. *Dr. Mellár Tamás*, a Központi Statisztikai Hivatal elnöke *Dancs Kálmánt*, a Fejér Megyei Igazgatóság osztályvezetőjét; *Mátrai Esztert*, a Bács-Kiskun Megyei Igazgatóság osztályvezetőjét; *Pintérmé Grósz Dojnat*, a Tolna Megyei Igazgatóság osztályvezetőjét és *dr. Falussy Bélát*, a Társadalomstatistikai főosztály főtanácsosát az Életmód–Időmérleg felvétel végrehajtásában végzett munkájuk elismeréseként; *Deák Istvánt*, a Heves Megyei Igazgatóság igazgatóhelyettesét, *Illyés Andrást*, a Somogy Megyei Igazgatóság osztályvezetőjét, *Jónás Istvánt*, a Baranya Megyei Igazgatóság osztályvezetőjét, *Szabó Istvánt*, a Békés Megyei Igazgatóság osztályvezetőjét, *Szabó Tibort*, a Bor-

sod–Abaúj-Zemplén Megyei Igazgatóság osztályvezetőjét, *Tóth Lászlónét*, a Jász-Nagykun-Szolnok Megyei Igazgatóság főmunkatársát, *Gratzl Ferencet*, a Népszámlálási főosztály osztályvezetőjét, *dr. Lakatos Miklóst*, a Népszámlálási főosztály vezető-helyettesét, *Nagy Orbánt*, a Népszámlálási főosztály szakfőtanácsosát, *Szűcs Zoltánt*, a Népszámlálási főosztály osztályvezetőjét, a 2001. évi népszámlálás szakmai előkészítésében és végrehajtásában végzett munkájuk elismeréseként; *Kátayné Csincsák Évát*, a Budapesti és Pest Megyei Igazgatóság osztályvezetőjét, mindkét feladat végrehajtásában végzett munkája elismeréseként *elnöki dicséretben* részesítette.

SZERVEZETI HÍREK – KÖZLEMÉNYEK

A hivatalos statisztika minőségéről rendezett nemzetközi konferenciát az EUROSTAT LEG (Leadership Expert Group on Quality) munkacsoportja Stockholmban, 2001. május 14-15-én. A Központi Statisztikai Hivatalt *Éltető Ödön*, a KSH ny. főosztályvezető-helyettese, *Mihályffy László*, a KSH tanácsadója és *Pukli Péter*, a KSH főtanácsadója képviselték. Két nap alatt 39 szekcióban 124 előadást kísérelhetett figyelemmel a 37 országból érkező 340 résztvevő. Az előadások témái érintették a teljes statisztikai adat-előállítási folyamatot a regiszterektől a kérdőívtervezés, a mintavétel és -becslés, a felhasználói igény felmérése, az interaktív adatelőkészítés, az automatikus kódolási eljárás, az imputálási technikák és a nemválaszolások témakörein át, egészen a metaadatbázis, a rövid távú jelzőszámok és a tájékoztatási módok kérdéséig. A zárójelentés ajánlásokat tartalmaz a „minőség” értelmezéséről, a minőség rendszerszerű szisztematikus megközelítéséről, valamint a statisztika „felhasználó-központúságának” növeléséről.

Konferencia a társadalmi struktúra változásairól. A strukturális változásokról és a modernizá-

ció kihívásairól a közép-kelet-európai országokban címmel a KSH Népeségtudományi Kutatóintézete, a Cseh Tudományos Akadémia Szociológiai Intézete és a berlini Társadalomkutató Tudományos Központ, valamint a lengyel és a szlovák Szociológiai Intézet 2001. május 11. és 13. között konferenciát rendeztek a posztoszocialista kelet-közép-európai országok strukturális változásairól és a modernizáció időszzerű kérdéseiről. Magyarországot *Spéder Zsolt*, a KSH Népeségtudományi Kutatóintézet igazgatója és *Harcsa István* a KSH főosztályvezetője képviselték. Valamennyi résztvevő ország bemutatta országjelentését. A magyar dolgozatot *Spéder Zsolt*, *Elekes Zsuzsa* (Budapesti Közgazdaságtudományi és Államigazgatási Egyetem), *Harcsa István* és *Róbert Péter* (Eötvös Loránd Tudományegyetem) készítettek el A magyarországi átalakulás körvonalai (The outlines of the transformation in Hungary) címmel.

A Nemzetközi Szociológiai Társaság Rétegződés–Mobilitás Szakosztálya (ISA RC28) 2001. április 25. és 28. között konferenciát tartott Mannheimben. A konferencia témái a következők voltak:

az iskolázási esélyek változása a különböző országokban; a társadalmi mobilitás legújabb trendjei; a jövedelmi egyenlőtlenségek nemzetközi összehasonlításban; a vállalkozóvá válást meghatározó különböző jellemzők; az iskolából a munkaerőpiacra történő átmenet legfontosabb jellemzői a különböző országokban; a felnőtté válás különböző szempontjai. A Központi Statisztikai Hivatal *Bukodi Erzsébet*, a Társadalomstatistikai főosztály osztályvezetője képviselte, aki a vállalkozóvá válással foglalkozó szekció munkájában vett részt és az önállóvá válás legfontosabb magyarországi jellemzőiről tartott előadást. Ebben a témában az elmúlt évben egy nemzetközi összehasonlító kutatás indult. A kutatási programban több mint tíz európai ország képviselteti magát. A konferencián előadások hangzóttak el a társadalmi mobilitással, az iskolázási egyenlőtlenségekkel és a felnőtté válás jellemzőivel foglalkozó munkacsoportokban is, melyek jó kiindulásul és például szolgálhatnak a hasonló tárgyú hazai elemzések elkészítéséhez.

Az oktatás szerepe a gazdaság és a társadalom fejlődésében címmel megjelent kiadvány azt vizsgálja, hogy nemzetközi összehasonlításban milyen tapasztalatok érvényesültek az oktatási célra fordított kiadások és a gazdasági növekedés összefüggésének terén a XX. század utolsó évtizedében és melyek a jelenlegi magyarországi helyzet fő jellemzői. Az egyes fejezetek bemutatják az iskolai végzettség szerepét a fenntartható növekedésben; az oktatásra fordított kiadások színvonalát; a népesség összetételét a legmagasabb iskolai végzettség szerint, az iskolai évek várható időtartamát az ezredfordulón; a felsőfokú végzettségük arányát és szerepét a gazdaság növekedésében; a speciális továbbképzést és a korszerű eszközök alkalmazását az iskolákban; a keresetek színvonalának alakulását iskolai végzettség szerint, végezetül Magyarország oktatási helyzetét az ezredfordulón.

(Az oktatás szerepe a gazdaság és a társadalom fejlődésében. [Összeállította: *Nyitrai Ferencné dr.*] Központi Statisztikai Hivatal. Budapest 2001. 75 old.)

Nonprofit szervezetek Magyarországon, 1999. A kiadvány – tartalmát és szerkezetét tekintve – beleillik abba a sorozatban, amely 1992 óta nyomon követi a nonprofit szektor magyarországi fejlődését. Az adatfeldolgozás során alkalmazott ún. rep-

rezentációs módszer egyrészt biztosította a korábbi eredményekkel való összehasonlíthatóságot, másrészt lehetőséget teremtett regionális, megyei szintű elemzésekre is. A módszertan részletes ismertetését a melléklet 1. fejezete tartalmazza. A korábbiakhoz képest ez a kiadvány jóval nagyobb számú nem elemzett táblaanyagot tartalmaz és ezáltal adattárként is használható.

(Nonprofit szervezetek Magyarországon 1999. Társadalomstatistikai Közlemények. Központi Statisztikai Hivatal. Budapest. 2001. 234 old.)

Helységnévkönyv 2001. január 1. Megjelent a Magyar Köztársaság Helységnévkönyve, melyet a Központi Statisztikai Hivatal 1992 óta évente, a területszervezési változások adataival bővítve tesz közzé nyolc fejezetben, a területszervezési változásokat tartalmazó Függelékkel. A fejezetek címei: Összefoglaló adatok; A helységek részletes adatai; Főbb területszervezési változások (1990 január 2. és 2001. január 1. között); A KSH által kibocsátott településazonosító törzsszámok; A körjegyzőségek jegyzéke megyénként; A kisebbségi önkormányzatok jegyzéke megyénként; A statisztikai kistérségekhez tartozó helységek jegyzéke megyénként; A helységek betűrendes névsora és a Függelék.

(A Magyar Köztársaság Helységnévkönyve. 2001. január 1. Központi Statisztikai Hivatal. Budapest. 2001. 214 old.)

A KSH Népeştudományi Kutatóintézet kutatási jelentései 67. kötete a családdal kapcsolatos magyarországi 1980 és 1998 közötti jogszabályok gyűjteményét tartalmazza, *Tárkányi Ákos* összeállításában. A szerző bevezető szavai szerint e jogszabálygyűjtemény készítésével az volt a célja, hogy megkönnyítse a családok problémáival foglalkozók munkáját. A jogszabályokat témakörönként, kronologikus sorrendben állította össze. A jogszabálygyűjtemény három forrása: a Törvények és rendeletek hivatalos gyűjteménye, 1980-1988-ig, egy CD-jogtár (1998-ra) és egy JURIX nevű CD-jogszabálygyűjtemény, amely az 1993. március 31-én hatályos vagy azóta keletkezett jogszabályokat tartalmazza.

(*Tárkányi Ákos*: A családdal kapcsolatos jogszabályok Magyarországon 1980-98-ig. Központi Statisztikai Hivatal Népeştudományi Kutatóintézetének kutatási jelentései 67. Központi Statisztikai Hivatal Népeştudományi Kutatóintézet. Budapest. 2001/2. 180 old.)

STATISZTIKAI IRODALMI FIGYELŐ

KÜLFÖLDI STATISZTIKAI IRODALOM

A STATISZTIKA ÁLTALÁNOS ELMÉLETE ÉS MÓDSZERTANA

HAND, D. J.:

AZ ORVOSI DIAGNOSZTIKÁT SZOLGÁLÓ STATISZTIKAI MÓDSZEREK

(Measuring diagnostic accuracy of statistical prediction rules.) – *Statistica Neerlandica*, 2001. 1. sz. 3–16. p.

Az orvosi diagnosztikát szolgáló statisztikai módszerek kidolgozása Fisher klasszikus lineáris diszkriminancia elemzéséig nyúlik vissza. Újabban a számítástechnikai eszközök fejlődése járult hozzá az osztályozás módszertanának és a mögötte álló elméleti kérdéseknek jobb megértéséhez. A diagnosztikai statisztikai módszerek két csoportba oszthatók aszerint, hogy diagnosztikai vagy mintavételi ismeretekre, támaszkodnak-e mint elméleti háttérre. Az előbbire példa a logisztikus diszkriminanciaelemzés és az osztályozási fa módszerek (rekurzív felosztások). Az ilyen módszerek kettőnél több osztályba sorolásra is felhasználhatók, de jelen esetben csak két osztályról lehet szó: beteg vagy nem beteg a vizsgált személy. Mintavételi paradigmára épült a klasszikus lineáris diszkriminanciaelemzés, ami ugyancsak kiterjeszhető kettőnél több osztály esetére. Az így kapott eredmény megbízhatóságát rontja, hogy általában kis mintákkal kénytelenek dolgozni, amit strukturált vagy simított kovariancia-mátrix felhasználásával próbáltak áthidalni. Mindkét csoportban nemparaméteres módszerek is kidolgozásra kerültek, amelyek helyi vagy relatív valószínűségi becsléseket használnak valamely osztályba tartozás megállapítására.

A statisztikai következtetés megbízhatóságára vonatkozó mérőszám megválasztása előtt általános kérdéseket is tisztázni kell. Ezek részben a diagnózis

készítésének céljával függnek össze: egyetlen betegre vagy egy beteg populációra kívánunk-e valamit megállapítani. Egyetlen új beteg esetén, annak valamely (betegségi) osztályba sorolása nem érinti a többi osztályba sorolást. Egész populáció osztályozásánál, mint amilyen adott betegségekre vonatkozóan végzett szűrővizsgálat, az egyedek valamely (remélhetőleg kisebb) csoportjának meghatározott betegségi osztályba sorolása a cél, és ilyenkor rendszerint kis valószínűségi határt kell előírni az odatarozást illetően.

Az is eldöntésre vár, hogy viszonylagos vagy abszolút megbízhatóságot várunk-e el a módszertől. Az előbbi többféle módszer összevetésekor jöhet szóba, az utóbbi az első vagy másodfajú hibára megállapított hibahatárra vonatkozik. Mind a viszonylagos, mind az abszolút megbízhatóságnál különbséget kell tenni az érdemi különbözőség és a szignifikáns különbség között. Az érdemi különbözőség esetenként „gyakorlati” kérdésekre vonatkozik, mint amilyen a diagnosztikai módszer megváltoztatásának a költségkihatása.

A szerző a két osztályba sorolás kérdéseit tárgyalja részletesen, és különböző betegségekre vonatkozó példákat ismertet, illetve helyzeteket elemmez. A módszerek hatékonyságát a hibás döntések következményeinek értékelésével kívánja felmérni, ami viszont nem mentes a szubjektív értékelésektől. Megoldás lehet, hogy nem írják elő azt a valószínűségi határt, amelyen az osztályba sorolás múlik, hanem a valószínűség szélesebb skálájára végzik el az értékelést. Az így kapott görbe alatti terület a módszer hatékonyságának jellemzésére szolgálhat: egyenértékű a Gini-indexszel vagy a Mann-Whitney-Wilcoxon-féle próbával.

Megjegyzés. A *Statisztikai Irodalmi Figyelő* rovatot a Központi Statisztikai Hivatal Könyvtár és Dokumentációs Szolgálat állítja össze. A rovat minden hónapban *Külföldi Statisztikai Irodalom* fejezetet (külföldi statisztikai és demográfiai könyvek és cikkek ismertetését), páratlan hónapban *Bibliográfiát* (a könyveket az MSZ 3423/2–84, az időszak kiadványokat az MSZ 3424/2–82 szabvány szerinti feldolgozásban), páros hónapokban *Külföldi folyóiratszemlét* tartalmaz. A Külföldi Statisztikai Irodalom fejezetet *Filó János* szerkesztette.

A statisztikai következtetés diagnosztikai megbízhatóságát általában olyan személyeken próbálják ki, akiről tudják, hogy a kérdéses betegségben szenvednek. A statisztikai módszer érthetően jól vizsgáljuk, viszont az is elképzelhető, hogy ez a minta valamely véletlen tulajdonságával van összefüggésben. A probléma megoldására kifinomult módszereket is igénybe vettek: keresztértékelés, bootstrap, jackknife becslés. A modell kiválasztásánál olyan stratégiát is alkalmaznak, ami egyszerre értékeli az illeszkedés mértékét és a modell bonyolultságát (utóbbit kedvezőtlennek tekinti), hogy kiszűrhetőek legyenek a túlságosan is jól illeszkedő módszerek. A rendelkezésre álló, illetve az eddig kipróbált

eljárásokat a szerző elegendőnek látja, és nem várja, hogy alapvetően új ötlettel álljanak elő, ami drámai módon javítaná az eredményeket. Arra viszont figyelni kell, hogy a mintasokaság egyéb tulajdonságai akár gyorsan is megváltozhatnak, ami ronthatja a használt módszer megbízhatóságát.

Végül az eljárás megbízhatósága függ attól is, hogy ki alkalmazza azt. A használatában jártas kutató kezében a módszer általában jobban vizsgáljuk, mint amikor az is szempont, miként lehet értelmezni a kapott eredményt a statisztikában nem jártas egészségügyi személyzet számára is érthető módon.

(Ism.: Szász Kálmán)

GAZDASÁGSTATISZTIKA

ÁRINDEXKUTATÁSOK A JÖVŐ ÉVTIZEDEKBEN

(Price index research in the coming decades.) –
Monthly Labor Review, 2000. 9. sz. 31–36. p.

A National Bureau of Economic Research (NBER) kutatási programja keretében, 1999 júliusában az árstatistika neves szakértői jöttek össze, hogy kifejtsek véleményüket arról, melyek azok a legfontosabbnak tartott problémák, amelyekkel az elkövetkező két évtizedben foglalkozni kell.

A konferencia legfontosabb megállapításait, valamint a széles körű, sok elvi és gyakorlati kérdést érintő vita lényegét *Ernst Berndt* professzor foglalta össze. Mondanivalója a következő négy kérdés köré csoportosítható.

1. Az árindexeket igen széleskörűen sokféle célra használják. Mely koncepciók irányába kell az indexeket fejleszteni és melyek az alkalmazás korlátai?

2. Az említett koncepcionális keretek között melyek azok a statisztikai kérdések, amelyekkel foglalkozni kell?

3. A statisztika keretei között melyek a megoldandó problémák?

4. Az egyes általános jellegű problémák megoldásainak milyen szektorspecifikus sajátosságai vannak?

Az Egyesült Államokban, a XX. század során, a Bureau of Labor Statistics (BLS) koncepciója, bár nem teljesen egyértelműen, a fogyasztói árindexet (CPI) a megélhetési költségek változása (COLI) irányából közelítette meg. Sok más országban, beleértve az EU koncepcióját is, a megközelítés iránya nem a COLI, mivel egyesek szerint ez nem jó közelítése az infláció mérőszámának. A Boskin jelentés azt javasolta, hogy a BLS a fogyasztói árak változását a megélhetési költségek irányából mérje. Választ kell viszont adni arra,

hogy bizonyos tartalmi kérdések, valamint a minőségi korrekciók miként értelmezhetők.

Nyilvánvaló, hogy a különböző felhasználóknak különböző árindexre van szükségük. Más tartalmú indexet szeretne egy bankár, mást a társadalombiztosítás stb. A BLS sokféle árindex kiszámítására képes, de a vitában arra is figyelmeztettek, hogy az árindexek különböző értelmezése, használata zavarhoz is vezethet, sőt politikai következményekkel is járhat. Világosan kell látni, hogy a különböző célra kiszámított árindexek felhasználása természetesen korlátozott.

A megélhetési költségek változása mérésének korlátai mind statisztikai, mind koncepcionális szempontból vita tárgyát képezték. Miként kell értelmezni a COLI keretei között az „életszínvonalat”, a „fogyasztói jólétet”? Hogyan lehet annak változását (változatlanágát) értelmezni? Valójában nagyon sok mindenről lehet szó, például a gazdasági, a környezeti, a klimatikus változásokról, vagy a bűnözésről, a zsúfoltságról stb. Egy lehetséges definíció megfogalmazható úgy, hogy mibe kerül a háztartásoknak az, hogy életkörülményeik fenntartása érdekében magasabb árakat kell fizetni.

Fontos és nehéz kérdés, hogy a termékek és fogyasztók milyen keretei között mérjük az árakat. Fontos, hogy meglegyen a kapcsolat az árindexek és a GDP között. Kérdés például az autók légszennyezése csökkentésének költségei vagy az egészségügyi kiadások miként számolhatók el az árindexek, illetve a nemzetijövedelem-számítások keretei között.

A koncepcionális kérdések között került még megemlítésre a fogyasztók viselkedésének a hasznosság oldaláról való megközelítése, azaz az előnyök és hátrányok összevetése, ami elméleti megközelítésben is nagyon bonyolult.

Végül a saját termelésű, eladásra szánt termékek és szolgáltatások számbavételének problémáit említi az összeállítás. A megélhetési költségek szempontjából el kell választani a tevékenységek input (anyagok, szolgáltatások) oldalát az outputtól.

Akármiilyen szűk, vagy tág keretek között is értelmezzük a COLI-t, a statisztikai mérések száma komoly probléma megoldását teszi szükségessé.

Hogyan lehet a több szempontból heterogén fogyasztók keresleti függvényeit aggregálni, hogyan lehet egy közös árindekshez eljutni. Ez felveti a „plutokrata”, vagy „demokrata” árindeks problémáját, illetve azt, hogy kell-e a lakosság különböző csoportjaira árindeksét számítani. A gyakorlat azt jelzi, hogy a különböző súlyozású indexek nem jeleznek számottevő eltérést. Inkább az a probléma, hogy a szegények és a gazdagok, a fiatalok és az öregek eltérő márkákat, választékot vásárolnak a különböző boltokban.

A konferencia résztvevői szerint az elkövetkező évtizedek kutatásait e területen a következő témákra kell irányítani.

1. Az a jelenlegi feltételezés, hogy a háztartások egy-egy csoportja ugyanazon termékért ugyanazt az árat fizeti, nyilvánvalóan nem igaz.

2. A jelenlegi felfogás szerint egy csoportra vonatkozó COLI mögött ugyanaz a népesség van mind a bázis-, mind a beszámolási időszakban. Ez a feltételezés azonban a különböző demográfiai változások miatt nem helytálló.

3. A COLI azt is feltételezi, hogy két időszak között az ízlések, szokások nem változnak, ami a reklám, az életkor változása miatt nem igaz.

4. A jelenlegi gyakorlat azt is feltételezi, hogy a környezeti változók (időjárás, légszennyezettség, helyi közlekedés stb.) azonosak. Ezek változását is kezelni kellene a COLI elméletében.

Számolni kell a szezonális ingadozásokkal, s kezelni kell valami módon azokat. Jellemző *Ralph Turvey* 1979. évi felmérésének eredménye, ami azt jelezte, hogy a különböző intézmények nagyon eltérő gyakorlatot folytatnak. Ez nem meglepő, hiszen a dolog természete nagy összehasonlíthatósági problémákat okoz.

Ezt a problémát úgy is meg lehet közelíteni, hogy három árindeksét számítsunk:

- a) havi indexeket hónapról hónapra,
- b) havi indexeket az előző év azonos hónapjához viszonyítva,
- c) év/év árindekszeteket, amit össze lehet vetni a 12 havi árindekszet mozgó átlagával. Az ilyen indexek kombinált használata lehet, hogy feleslegessé tenné a szezonális kiigazítások szokásos eljárását.

A minőségi korrekciók (quality adjustment) módszerei is folyamatosan viták tárgya. A BLS a minőségi változásokat leginkább a „hedonikus” módszerrel kezeli, melynek alkalmazása várhatóan

tovább bővül. Ebbe a körbe tartozik a lakás és lakás-felszerelés, televízió és a számítógépek indexe. Folyamatban van e módszer alkalmazásának előkészítése a telefonkészülékek, videó- és DVD-lejátszók, a hűtőgépek, mosó- és szárítógépek, mikrohullámú és audiokészülékek csoportjaira.

A hedonikus megközelítés használata leginkább gyakorlati problémák megoldását igényli. Végül is az a cél, hogy a termékek cserélődésének előnyei és hátrányai hogyan összegezhetők a fogyasztó szempontjából. Ez a terület az elkövetkező 10-15 év alatt sokat fejlődhet.

Jelentős fejlődését hozhat a szkennelt adatok használata, valamint az elektronikus adatgyűjtés. Ehhez szükség van a vállalatok közreműködésére, ami esetleg pótolhatja némely esetben a hagyományos árfelírást.

Számolni kell az internettel is. Már hozzáférhető az interneten olyan adatok, amelyek azt jelzik, hogy bizonyos termékek milyen (legolcsóbb) áron vásárolhatók meg.

Az árak megfigyelése, mérése termékcsoportonként (szektoronként) nagyon sajátos problémák megoldását igényli. Ezek között talán legnehezebb a szolgáltatások áralakulásának számbavétele. A probléma már a szolgáltatások outputjainak definícióival kezdődik, ami sok esetben nem (lehet) egységes. Két vásárló gyakorta nem ugyanazt a terméket „fogyasztja”. Igaz ez az építőipari tevékenységek esetében is. Erre a problémára nem lehet általános megoldást találni. Minden egyes eset külön-külön vizsgálatot igényel. Meg kell még említeni az egészségügyi szolgáltatásokat, melyek árváltozását talán a legnehezebb mérni. A statisztikusoknak folyamatosan szembe kell nézni ezekkel a problémákkal.

Végül hangsúlyt kapott az is, hogy a kutatások jelentős költségekkel járnak. Mérlegelni kell a különböző problémák kezelhetőségét, munkaigényét. De nagyon fontos az is, hogy a koncepcionális kérdéseket ne csak a költségek oldaláról vizsgáljuk, hanem azt is, hogy azok a felhasználók számára is érthetőek, elfogadhatók legyenek.

(Ism.: *Marton Ádám*)

ÚJ MUNKAERŐ-PIACI TRENDKÉK A KÖZÉP-EURÓPAI ORSZÁGOKBAN

(Recent labour market trends in the Central European Countries.) – *Central European Countries Employment and Labour Market Review*, 2000. 2. sz. 7–11.p.

A tanulmány 12 ország (Albánia, Bulgária, Csehország, Észtország, Lengyelország, Lettország,

Litvánia, Macedónia, Magyarország, Románia, Szlovákia és Szlovénia) munkaerő-piaci helyzetét vizsgálja.

A gazdasági teljesítmények 1998 utolsó hónapjaiban megfigyelt csökkenése 1999 első felében folytatódott, sőt kifejezettebbé vált a legtöbb átmeneti országban. Szlovénia az egyetlen kivétel, ahol a gazdasági növekedés ebben az időszakban nagyobb volt, mint egy évvel korábban. 1998 második felében a GDP növekedési üteme relatíve gyors maradt Magyarországon, bár gyorsuló ütemét elveszítette. Hét év erős gazdasági teljesítménye után Lengyelország a lehajló növekedési periódusba lépett ugyan ebben az időszakban. Szlovákia teljesítménye is messze elmaradt a megelőző évek színvonalától. Romániában három egymást követő évben is csökkent a GDP. Csehországban 1999 második negyedévében lassú növekedés indult meg, az 1998. év elején beinduló recesszió után. Magyarország kivételével az említett országokban esett az ipari termelés, ami legnagyobb mértékben Bulgáriát, Romániát, Szlovákiát és Csehországot érintette.

A relatíve nagy gazdasági növekedés éveit követően a Balti-államok gazdaságára különösen súlyos hatással volt az orosz válság. 1998 negyedik negyedévével kezdve a gazdasági teljesítmények gyorsan csökkentek. A negatív trendek 1999 első felében folytatódtak: a GDP Litvániában 4,8 százalékkal, Észtországban 3,9 százalékkal, Lettországon 2,0 százalékkal visszaesett. A GDP csökkenése mindhárom országban együtt járt az ipari termelés mérséklődésével.

Az átmeneti országok gazdasági teljesítményének esése külső okokkal magyarázható: elsősorban a nemzetközi pénzügyi válsággal, az orosz krízissel és a koszovói konfliktussal, melynek eredményeként a GDP növekedési üteme messze elmaradt az EU átlagától. A visszaesésben hazai okok is szerepet játszottak: az intézményi reformok lassúsága, a vállalatok átalakulása, a növekvő hitelkamatok és adósságtérhek, valamint a banki és a vállalati szektor tehetetlensége. Mindezek gyengítették a keresletet az EU piacain, amelyek az ipari termelés, az import és a GDP visszaesésével jártak. A koszovói konfliktus negatív következményei különösen erősek voltak Délkelet-Európában, főleg Bulgáriában, Romániában és Macedóniában a közlekedési útvonalak szétszakítása miatt.

1999 első félévében az infláció a legtöbb átmeneti országban lassú ütemben, folyamatosan csökkent, sőt Bulgáriában defláció volt. Romániában is csökkent a fogyasztói árindex, de éves szinten ugyanebben az időszakban még mindig kétszámjegyű volt. 1999 első felében az export figyelemre méltó ütemben csökkent a legtöbb országban, kivéve

Magyarországot, ahol az export mintegy 6 százalékos növekedést mutatott (euróban mérve), körülbelül annyit, mint a megelőző évben. Románia és Szlovákia kivételével a fizetési mérleg egyensúlya a régió legtöbb országában romlott. Szlovénia a függetlenség kivívása óta először nézett szembe komoly deflációval az import növelésének és az áfa 1999. évközepe bevezetésének következményeként.

A foglalkoztatás növekedési ütemének mérséklődése 1998-ban folytatódott. A rendelkezésre álló munkaerőfelvétel-adatok szerint 1999 első hónapjaiban a foglalkoztatás nyolc közép-európai országban csökkent, legfigyelemreméltóbb mértékben Bulgáriában (5,7 százalékpont). Nincsenek adatok Lettországról, Macedóniáról és Albániáról. A foglalkoztatás szintje csak Magyarországon és Litvániában nőtt, egyenként mintegy 3,5 százalékponttal. Mindkét országban, de különösen Litvániában a női foglalkoztatás mértéke erősebben nőtt, mint a férfiaké. Észtországban, Lengyelországban és Szlovákiában a foglalkoztatás csökkenése a nőknél kevésbé volt jellemző, mint a férfiaknál. Bulgáriában, Szlovéniában és Csehországban a férfi és a női foglalkoztatás szintje csaknem hasonló mértékben mérséklődött, míg Romániában a csökkenés teljes mértékben a férfiakat érintette.

A foglalkoztatás idősorának a tanulmányban közzétett kezdőpontjától, 1996-tól eltelt teljes periódus alatt a foglalkoztatási ráta erősen csökkent Csehországban, Romániában és Szlovákiában. Magyarországon és Litvániában a foglalkoztatás szintjének csökkenése 1998-ban is folytatódott, ami tovább tartott 1999 első hónapjaiban is. Lengyelország és Szlovénia az 1996 és 1998 közötti időszakban növelni tudta foglalkoztatottainak számát, ami 1999-ben csökkenésbe csapott át. Lettországon és Macedóniában 1998-ban nőtt a foglalkoztatás szintje, de 1999 első félévéről nem állnak rendelkezésre adatok.

A közép-európai és az EU-beli országok foglalkoztatási rátáinak összehasonlítása 1998-ra vonatkozóan azt mutatja, hogy az előbbieket átlaga (62 százalék) éppen csak meghaladja az EU országainak átlagát (61%). A foglalkoztatási ráta a romániai 70 százalékos aránytól az 54 százalékos arányt mutató Magyarorszáig terjed, eltekintve az extrém adattal rendelkező Macedóniától. A foglalkoztatásban a nemek közötti különbség a közép-európai országokban kisebb, mint az Európai Unió legtöbb országában. A női foglalkoztatási ráták a legmagasabbak Észtországban, Romániában, Csehországban és Litvániában (körülbelül 60 százalék) míg a legkisebb Macedóniában (30%) és Albániában (39%). A két utóbbi ország, valamint Bulgária és Magyarország kivételével a másik nyolc átmeneti országban a női

foglalkoztatás szintje magasabb, mint az Európai Unióban.

A terveződésből a piacgazdaságba való átmenet lényeges változásokat hozott a gazdasági szerkezetben és ebből következően a foglalkoztatási összetételben is. A legtöbb országban csökkent az iparban és a mezőgazdaságban foglalkoztatottak aránya és erőteljesen nőtt a szolgáltatási szektorban dolgozóké.

A foglalkoztatottak arányainak változása a GDP-ben is hasonló módosulást eredményezett. Eltekintve Albánia extrém esetétől, ahol a foglalkoztatottak 70 százaléka még ma is a mezőgazdaságban dolgozik, Románia a meggyőzőbb példa arra, hogy sokan az iparból és az építőiparból a mezőgazdaságban találtak foglalkoztatást. A mezőgazdaságban foglalkoztatottak aránya a teljes átmeneti periódusban növekedett, amely 1998-ban Romániában elérte a 40 százalékot. Hasonló trendek figyelhetők meg Bulgáriában és Litvániában, ahol ezek az arányok 26, illetve 21 százalékot tesznek ki. E folyamatok a nagy állami tulajdonú mezőgazdasági kombinátók, üzemek privatizációjával és az ingyenes földtulajdonhoz juttatással (Litvánia) vannak összefüggésben. Emellett az emberek a munkanélküliségtől való félelmükben is gyakran a mezőgazdasági szektorban kerestek jövedelmi lehetőséget.

Az iparban a dolgozók hatalmas állásvesztései ellenére a foglalkoztatottak aránya még ma is magasabb mint az EU átlaga. A legtöbb országban (Bulgária, Észtország, Magyarország, Lengyelország és Románia) az ipari foglalkoztatottak aránya mintegy egyharmada az összes foglalkoztatottnak. Az egy főre jutó GDP alapján az átmeneti országok két legfejlettebbike Szlovénia, Csehország, valamint Szlovákia mutatja a legmagasabb ipari foglalkoztatottsági arányt, elérve a 40 százalékot is. A skála másik végén Albánia áll, ahol ugyanez az arány 9 százalék, majd Lettország, Litvánia és Macedónia következik az ipari foglalkoztatottak egyaránt 27 százalékaival. Az Európai Unió országaiban ez a ráta 29,5 százalék.

Románia kivételével 1993 óta a szolgáltatási szektor foglalkoztatottjainak aránya mindegyik átmeneti országban nőtt. Ebben az országcsoportban, ahol az átalakulás már az 1980-as évek végén megkezdődött, Magyarországon a legmagasabb ez az arány (majdnem 60 százalék). Lengyelországban és Bulgáriában a terciér szektor aránya meghaladta a 40 százalékot, Romániában nem érte el az egyharmados arányt. Csehországban és Szlovákiában az e szektorban foglalkoztatottak aránya alig haladta meg az 50 százalékot, míg Szlovéniában valamelyest elmaradt ettől az értéktől. Minden átmeneti ország terciér

szektorbeli aránya elmaradt az EU átlagától (65,5 százalék 1998-ban), de Magyarország és Észtország adata meghaladta a déli EU-beli országok (Görögország, Portugália, Spanyolország) átlagát.

A szolgáltatási szektor a legfőbb foglalkoztatója a nőknek, az átmeneti országokban is. A nők aránya ebben a szektorban mindenütt magasabb volt (kivéve Romániát), mint a férfiaké, mely arányok 71 százalék (Csehország, Észtország, Magyarország és Szlovákia) és 60 százalék (Litvánia, Lengyelország, Szlovénia) között szóródtak. Lettországon pontosan 65 százalékat adták a nők a szolgáltatási szektor dolgozóinak. Az 1999. évi adatok e szektor erősödését jelzik Magyarországon, Szlovéniában és Szlovákiában. Romániában a foglalkoztatási trendek a korábbi évek folytatódását mutatják. A gazdasági visszaesés tükröződik a munkanélküliség növekedésében is a régió csaknem minden országában. A munkaerő-felvétel 1999. évi adatai a munkanélküliség emelkedését jelzik a közép-európai országokban, kivéve Magyarországot és Litvániát. Szlovéniában a ráta az előző évi szinten maradt. A munkanélküliségi ráta 6,2 százalék (Románia) és 15,8 százalék (Szlovákia) között szóródott. A két utóbbi országban a munkanélküliségi ráta egy év alatt 4 százalékponttal emelkedett. Bulgáriában, Csehországban és Észtországban ugyanez 2 százalékpont volt. Lengyelországban, ahol a munkanélküliségi ráta az elmúlt években folyamatosan csökkent, az orosz válság miatt szintén nőttek a foglalkoztatási feszültségek. Csak két ország, Magyarország és Litvánia tudott győztesen kikerülni a nehéz helyzetből, azaz csökkent a munkanélküliségi ráta, Magyarországon a férfiaké, míg Litvániában a nőké mérséklődött gyorsabban.

A munkaerő-felmérés adatai szerint 1999 második negyedévében a régióban az álláskeresők száma elérte az 5,6 milliót a régióban, mely 15 százalékkal több, mint egy évvel korábban. Csehországban a munkanélküliek száma az egy évvel korábbihoz viszonyítva 44 százalékkal nőtt, Szlovákiában 31, Észtországban 20 százalékkal. A regisztrált munkanélküliek aránya Észtországban volt a legkisebb (3,2%), Szlovákiában és Albániában a legmagasabb (majdnem 18 százalék). Mintegy 9 százalékos értéket jelentettek Lettországból, Magyarországról és Csehországból.

A női munkanélküliségi ráta 6 százalék (Magyarország és Románia) és 16 százalék között (Szlovákia) változott. Négy országban (Észtország, Magyarország, Litvánia, Románia) a női munkanélküliségi ráta alacsonyabb mint a férfiaké. Szlovákiában és Bulgáriában a férfiakéhoz hasonló.

A fiatalok (15-24 évesek) munkanélküliségi rátája jelentősen magasabb mind az átlagnál, mind az idősebbekre vonatkozó adatnál. A legtöbb országban a fiatalok munkanélküliségi rátája kétszer magasabb az átlagnál, sőt Romániában és Lettorszában háromszor annyi. Ez alól csak Észtország kivétel, ahol ez a különbség jóval kisebb. A legmagasabb ifjúsági munkanélküliségi rátát Lettországból (38%) és Bulgáriából (33%) jelentették és a legalacsonyabbat Csehországból (12%) és Magyarországról (13%). Mintegy 22-24 százalékos e ráta Litvániában, Lengyelországban és Szlovákiában. Megdöbbentően magas viszont (70%) Macedóniában. Az ifjúság munkanélküliségi rátája az EU-országok között is különösen nagy Olaszországban és Görögországban, ahol ez az arány mintegy 30 százalék. A rendelkezésre álló adatok szerint a munkanélküliségi ráta magasabbak a fiatal nőknél, mint a férfiaknál Csehországban, Macedóniában, Lengyelországban, Romániában és Szlovéniában, fordított a helyzet Bulgáriában, Észtországban és Magyarországon. Szlovákiában a nemek között e téren nincs lényeges különbség.

(Ism.: *Hajnal Béla*)

BRUVOLL, A. – NYBORG, K. – HALVORSEN, B.:

A HÁZTARTÁSI HULLADÉKOK
HÁZILAGOS SZÉTVÁLOGATÁSA

(Household sorting of waste at source.) – *Economic Survey*, 2000. 4. sz. 26–34. p.

A norvég kormány a legutóbbi években 2010-re elérendő célként irányozta elő, hogy a keletkező hulladékok legalább 75 százalékát újra dolgozzák fel, vagy hasznosítsák energia visszanyerésére. A széles körű másodlagos feldolgozás (recycling) fontos előfeltétele, hogy a háztartások a náluk keletkező hulladékokat felhasználási célok szerint osztályozzák.

A norvég háztartások már jelenleg is foglalkoznak a hulladékok szétválogatásával s az újrafelhasználható csomagolóanyagokat (mint például az összehajtogatott kartondobozokat vagy a kimosott konzervek üvegeket) sokan rendszeresen elviszik az utcákon felállított gyűjtőtartályokba. A norvég Statisztikai Hivatal adatai szerint 1992 és 1998 között az újrafeldolgozásra átadott háztartási hulladékok aránya 8 százalékról 33 százalékra emelkedett. A szelektív hulladékgyűjtés terjedése ellenére jelenleg is megoszlanak a vélemények e törekvések környezeti hasznosságáról, a felmerülő költségek nagyságáról és főleg arról, hogy a lakosság milyen mértékben hajlandó vállalni az ezzel járó pótlólagos költségeket.

A szerzők annak az adatgyűjtésnek a főbb eredményeit ismertetik, amelyet a norvég Statisztikai Hivatal egy ún. „omnibusz felvétel” keretében hajtott végre 1999. november és december havában, a hulladékhasznosulási információk megalapozása céljából. Az összetett tematikájú „omnibusz felvétel” számos háttérváltozóról (például kor, nem, családi állapot, jövedelem stb.) tartalmaz kérdéseket, amelyek köre alkalmanként speciális témákkal is kiegészülhet. A tárgyalt megfigyelés esetében a hulladékválogatással, a dohányzással és a kommunális munkákkal kapcsolatosan tettek fel pótlólagos kérdéseket az adatszolgáltatóknak. A mintavételi terv szerinti 2000 elemű minta a norvég népesség 16-79 éves korcsoportjaira vonatkozott. A válaszolási arány kissé alatta maradt a 60 százalékéknak, a következőkben bemutatott eredmények tehát összesen 1162 válaszoló adataira támaszkodnak. A norvég Statisztikai Hivatal szokásos mintavételi gyakorlatának megfelelően az adatgyűjtésnél kétféle eljárást alkalmaztak: a megkérdezettek 76 százalékát otthonukban keresték fel a kérdezőbiztosok, míg a többiekkel telefonon bonyolították le az interjút.

A megfigyelés keretében elsősorban azt kívánták megállapítani, hogy a norvég háztartásokban mennyi időt fordítanak a keletkező hulladékok szétválogatására, illetve megtisztítására, s hogy az utóbbi műveletek milyen energiafogyasztási többlettel járnak. Fontos vizsgálati témának tekintették e tevékenységek motivációs hátterének feltárását, megkülönböztetett figyelmet fordítva arra, hogy a háztartások mennyire tekintik terhesnek a hulladék-szortírozást s hajlandók lennének-e fizetni azért, hogy ezt – társadalmilag szervezett módon – mások végezzék el helyettük.

A válaszolók 93 százaléka nyilatkozott úgy, hogy legalább részben rendszeresen szétválogatja a hasznosítható háztartási hulladékokat, közelebbről a papírt és a hullámkarton-papírt, az üvegeket, az italok karton-csomagolását, a komposztálható ételmaradékokat, továbbá a fémeket, illetve a műanyagokat. Ez a sorrend egyébként a gyakoriságot is kifejezi, a kapott eredmények ugyanis egyértelműen jelezték, hogy a legelterjedtebb a papír és kartonpapír, illetve az üvegek kiválogatása és összegyűjtése (arányuk 60-70 százalékot képviselt). A szortírozás mértékének jellemzésére a következő négy válaszlehetőség szolgált: az összeset, a többséget, kismérszt, vagy semmit sem szortíroznak a kérdéses hulladékfajtából.

Megjegyzendő, hogy a megfigyelési eredmények viszonylag kedvezőbb képet mutatnak, mint a hulladékgyűjtés tényleges adatai. Az 1998. évben például, az itt jelzett 20 százalék körüli aránnyal szemben, csak a háztartások mintegy 10 százalékától

származott műanyag hulladék. Az eltérés egyik oka az lehet, hogy válaszadók az italok betétidijas csomagolóanyagainak visszaszolgáltatását is figyelembe vették, holott a megfigyelési előírások szerint ez nem számított bele a háztartási hulladék szortírozásának időszükségletébe.

Azok a válaszolók, akik a háztartásukban keletkező hulladékokból kiválogatják az újrahasznosítható anyagokat, hetenként átlagosan 30 percet fordítanak erre a tevékenységre. Ezen belül maga a szortírozás, a papírok összehajtogatása mintegy 14 percet, az üvegek kimosása 9 percet, míg a szállítás a gyűjtőkonténerekhez átlagban 7 percet vett igénybe. A megfigyelés csak a felnőttekre vonatkozott, s az érintett személy saját, egyéni tevékenységét tudakolták. Háztartásonként 1,7 felnőttel számolva, éves viszonylatban háztartásonként 41 órára tehető az ott keletkezett hulladékok szelektív kezelésének teljes időszükséglete.

A kormányzati célok nagyságrendi érzékeltetése céljából a teljes hulladékmennyiség alapulvételével is végeztek közelítő számításokat. A norvég Statisztikai Hivatal 2000. évi adatainak tanúsága szerint a háztartások összesen 452 ezer tonna hulladékot bocsátottak rendelkezésre másodlagos hasznosítási célokra. A reprezentatív megfigyelés eredményeit a teljes felnőtt népességre jellemzőnek tekintve tehát tonnánként mintegy 185 óra ráfordítást igényli, ha a hasznosítható hulladékok fajtankénti szelektálását a keletkezésük helyén, a háztartások hajtják végre.

A mintavételi eredmények elfogadhatóságának elbírálásánál két független információforrásra is támaszkodhattak. Az 1991–1992. évekre vonatkozó időmérleg szerint lakástakarításra, mosogatásra stb. személynként átlagosan heti 5 órát fordítottak. Nem tűnik valószínűnek, hogy e háztartási munkák teljes időszükségletének 5-10 százalékát a hulladékszortírozás vegye igénybe. Feltételezhetőbb, hogy a közvetlen rákérdezés a reálisnál némileg magasabb adatokat eredményezett.

Hasonló következtetés vonható le abból az összehasonlításból is, amelyet a Svéd Fogyasztók Érdekképviselőinek laboratóriumi feltételek között végzett vizsgálatára támaszkodva hajtottak végre. A svéd adatok szerint a csomagolási rendeltetésű hulladékok megtisztításának heti időszükséglete 22 percre tehető, míg a norvég mintavétel válasza szerint erre a tevékenységre átlagosan 14 percet fordítanak. Nem hagyható azonban figyelmen kívül, hogy a svéd vizsgálat 4 fős háztartásra vonatkozott, szemben a ténylegesen 2,2 fős norvég átlaggal. Végül soron tehát ez az összevetés is azt jelzi, hogy a kimutatott norvég időfelhasználásokat viszonylag magasnak kell tekinteni.

A svéd adatok segítségével egyébként kísérletet tettek annak becslésére is, hogy milyen energia-, illetve vízfelhasználási többlettel járhat, ha a háztartások melegvizet használnak a visszatérülő csomagolóanyagok tisztítására. A svéd adatok erre a célra heti 50 liter vízfogyasztást jeleztek. Norvég viszonylatban is hasonló arányokat feltételezve, éves szinten háztartásonként 3,1 köbméter melegvíz-többlettel lehet számolni. A víz legalább 35 Celsius fokra melegítéséhez szükséges villamos energia mennyiség mintegy 0,3 százalékkal növelné a norvég háztartások éves villamos energia fogyasztását.

A fenti feltételekkel becsült idő-, víz-, illetve energiaráfordításokat az összes szelektált háztartási hulladék mennyiségére vonatkoztatva, tonnánként évi 186 órányi idő-, 7,3 köbméter víz-, illetve 218 kWó villamosenergia-felhasználási igényre lehet számítani. A központi gyűjtőkonténerek megközelítése (az átlagosan 7 perces időszükségletet alapul véve és 10 kilométer távolságra 0,9 literes benzinfogyasztást feltételezve) tonnánként mintegy 140 korona szállítási költséget jelentene.

Az ismertetett eredmények bizonytalansági tényezőit mérlegelve a szerzők több módszertani problémára is felhívják a figyelmet. Véleményük szerint nem zárható ki annak lehetősége, hogy a válaszolók a valóságosnál pozitívabb magatartást jeleztek a szelektív hulladékgyűjtéssel kapcsolatban. Ez történhetett a kedvezőbb benyomás elérése érdekében, de azért is, mert a közvetlenül megkérdezettek önkéntlenül is hajlamosak lehetnek az egyéni teljesítmények túlbecslésére, főleg ha ilyen viszonylag bonyolult, sehol nem dokumentált adatokról kell nyilatkozniuk. A közölt időfelhasználások gyakran észlelt „kerekítése” is többnyire felfelé térítette el az eredményeket. Az átfogó számításoknál azt sem szabad figyelmen kívül hagyni, hogy a mintavételi körön kívüli népességre kedvezőtlenebb arányok lehetnek jellemzők.

A lakosság várható magatartásának megismerése érdekében a megfigyelés a szelektív hulladékkezelést végzők jelenlegi felfogásáról, tevékenységük motivációjáról is igyekezett információkat gyűjteni. Hat állítást fogalmaztak meg, és arra kérték a válaszolókat, jelöljék meg, hogy milyen mértékben fogadják el ezeket. (A lehetséges válaszok a következők voltak: egyetért, részben egyetért, részben nem ért egyet, nem ért egyet, nem tudja.)

A helyzet ismeretéhez hozzátartozik, hogy a környezetszennyezést korlátozó norvég törvény alapján az önkormányzatok bírságot is szabhatnak ki azokra a háztartásokra, amelyek nem tesznek eleget hulladékválogatási kötelezettségeiknek, a gyakorlatban azonban erre nem igen akad példa. Mindeneset-

re a válaszolók 63 százaléka hatóságilag elrendelt követelményként érzékeli ezeket a feladatokat. Sokan önkéntes alapon is hajlandóan mutatkoznak elvégzésükre, de nem támogatnák a válogatandó hulladékok körének lényeges kiterjesztését. Arra a kérdésre, hogy önmagában kellemes tevékenységnek tekinti-e a hulladékszelektálást 44 százalék adott ki-fejezetten, és 18 százalék részben tagadó választ.

A progresszív jellegű motivációk közül egyértelműen a környezetvédelmi megfontolások dominálnak: a válaszolók 97 százaléka ezt tekinti fő indoknak. A „jó cselekedet öröme” a válaszolók mintegy 40 százaléka döntésénél játszhat pozitív szerepet. A válaszolók 73 százaléka morális kötelességének tekinti a kormányzati célok támogatását ezen a fontos területen és mintegy 40 százalékuk semmi esetre sem szeretné kockáztatni a munkatöbblet miatt saját társadalmi megbecsülését rokonai és ismerősei részéről.

A megfigyelés kiemelkedő fontosságú kérdésének tekintették annak tisztázását, hogy amennyiben azonos környezetvédelmi és változatlan anyagi feltételek között, közösségileg szervezett módon a háztartáson kívüli körben oldanak meg a hasznosítható hulladékok szortírozását, mennyien tartanának igényt erre a szolgáltatásra. A válaszolók 72 százaléka adott igenlő választ a felvetésre és csak 27 százalék végezné továbbra is saját maga a válogatást. Az előbbieket közül sajátos módon azok lennének hajlandók viszonylag nagyobb (óránkénti) összeget is fizetni az ilyen szolgáltatásért, akik jelenleg is kevés időt fordítanak a hulladékok szelektálására. A társadalmilag indokolt költségek megállapítása a szerzők véleménye szerint még további, részletesebb információkat és számításokat igényel, a végső döntésnél azonban feltétlenül fontos szerephez kell jutniuk a szociális megfontolásoknak is.

(Ism.: *Tűű Lászlóné*)

SMIRNOV, S. N.:

A REGISZTRÁLT MUNKANÉLKÜLISÉG PROGNOSZTIZÁLÁSÁNAK MÓDSZERE

(Ispol'zovanie pokazatelej gosudarstvennoj statistiki dija prognozirovaniya urovnja registrirovannoj bezraboticy.) – *Voprosy Statistiki*, 2000. 7. sz. 7–13. p.

Az orosz foglalkoztatáspolitikai középtávú prognózisaihoz – melyekben az ország régióinak munkaerőpiacát mérik fel – olyan mutatók szükségesek, amelyek a regionálisan differenciált foglalkoztatáspolitikát tükrözik. A 90-es évek második felében ki-

dolgozott módszer régiócsoportosítási alapjául a következő szempontok szolgáltak.

1. A régió munkaerő-piaci jellemzői mutatóinak alapja a gazdaságilag aktív lakossághoz viszonyított regisztrált munkanélküliek száma. A mutató finomításához kidolgozták például a munkaerőpiac feszültségi mutatóját figyelembe véve az egy álláshirdetésre jelentkező regisztrált munkanélküliek számát és az egy évet meghaladó munkanélküliek arányát, mely jelzi a munkanélküliség állandósulását.

2. A gazdasági mutatók azért szükségesek, hogy az adott területi egységekben megállapíthassák, megfelel-e a foglalkoztatáspolitikai a gazdasági helyzetnek. Csoportosítást végeztek az iparcikktermelés dinamikája szerint; vizsgálták a gazdaságtalan iparvállalatok százalékos arányát is. A gazdasági szférát jellemzi az alaptőke növelési üteme. Két egymást követő év első félévének tőkeemelési mértéke szintén meghatározó a munkanélküliség tartóssá válásában. Mutatóként alkalmazták az egyes cégek, vállalatok teljes tőkeemeléséhez viszonyított saját tőkéjének arányát. Mindezek a munkaerőpiac javulásának előfutárai az egyes régiókban, következésképp ezeknek a mutatóknak mind a munkaadók, mind a foglalkoztatást elősegítő szervezetek foglalkoztatáspolitikai programjainak készítésében fontos szerepük van.

3. A szociális mutatók közül az elemzések során vizsgálták a lakosság reális jövedelemnövekedési ütemét. Itt is két év (1995–1996) első félévének adatait vetették össze. Több korrelációs elemzést végeztek, de olyan térséget nem találtak, ahol az alacsony szintű munkanélküliség találkozott volna a legalacsonyabb jövedelemnövekedési ütemmel. Vizsgálták továbbá a lakosság pénzintézeteknél történő megtakarításait is. A szociális mutatók főleg perspektivikus jellegű foglalkoztatáspolitikai döntéseknél meghatározók. Az összehasonlítható áron mért árukereskedelem növekedési ütemének mutatója szintén alkalmazható a csoportosításokban, de nem meghatározó jelleggel, inkább a regionális sajátosságok vizsgálatát segítve elő. A munkanélküliek szempontjából válságos területeken ugyanakkor tapasztalható a kiskereskedelmi áruforgalom növekedése. Ez utóbbi általában a nem adózó szektor foglalkozásnövekedését is jelzi. Alacsony munkanélküliség a kiskereskedelmi áruforgalom gyors csökkenésével együtt nem volt kimutatható.

4. A demográfiai mutatók sorából hosszú távú jelleggel felhasználták a népesség természetes mozgását. A területi mobilitás, a költözések szintén – adott esetben lényegesen – befolyásolják a foglalkoztatáspolitikai hatékonyságát, illetve jelenthetik a krízishelyzet megoldását is.

A foglalkoztatáspolitikai helyes megválasztásához számításba vett gazdasági, szociális és demográfiai mutatók az egyes régióval foglalkozó szakemberek számára lehetőséget nyújtanak a kulcskérdések megoldására.

A számítások az átlagra jellemző csoportot határoztak meg, összesítették a csoportok számait, és a kapott összeget elosztották az alkalmazott mutatók számával. A legjobb mutatóval rendelkezőket az egyes számú csoportba, a legrosszabbakat a 7. csoportba sorolták.

A javasolt módszer alkalmasnak bizonyult az Orosz Föderáció régióiban nyilvántartott munkanél-

külségi szint változásainak prognosztizálására. A kapott adatok azért is jól hasznosíthatók, mert a vizsgált időszak felöleli a válságos és a válságot követő időszakokat is. Az ajánlott módszer megszorításokkal bár, de alkalmazható a regisztrált munkanélkülségi szint változásainak hosszú távú prognosztizálására a föderáción kívüli régiókban is.

Az 1996–1999 közötti években csökkent a nyilvántartásba vett munkanélküliek száma: a várhatóan növekvő munkanélküliséggel prognosztizált 20 régióban átlagosan 42 százalékkal, a munkanélküliség nem várt növekedésébe tartozó 18 régiónál 50 száza-

lékkal. Mindemellett öt régióban (25%), ahol növekedést prognosztizáltak, nőtt a munkanélküliek száma, miközben, ahol nem prognosztizáltak, ott 11 százaléknit tett ki a növekedés. Ez arra utal, hogy a munkanélküliségi szint prognosztizáltan növekvő és nem növekvő régiói középszintjének változása között állandó eltérés mutatható ki.

Mindent egybevetve a tanulmány szerzője úgy véli, hogy a prognózis statisztikai bázisát másfél vagy két éves gyakorisággal célszerű megújítani.

(Ism.: *Bíró Júlia*)

TÁRSADALOMSTATISZTIKA – DEMOGRÁFIA

SCHULZ, R.:

A VILÁG NÉPESSÉGÉNEK ÖREGEDÉSE

(Die Alterung der Weltbevölkerung.) – *Zeitschrift für Bevölkerungswissenschaft*, 2000. 2. sz. 267–289. p.)

A népességöregedés problémájával ma túlnyomórészt csak a fejlett ipari országokban foglalkoznak, ugyanakkor a fejlődő országokat is érinti, hiszen itt is megfigyelhető a születések számának visszaesése, ami bizonyosan a lakosság öregedéséhez vezet majd. Ezekben az országokban a folyamat néhány évtized alatt megy végbe, míg a fejlett országokban száz évre nyúlik vissza.

Ma általánosan elfogadott álláspont, hogy a népesség demográfiai értelemben vett öregedését alapvetően a születések számának visszaesése határozza meg, és a halálzási arányszámok csökkenése erősíti. Legalább ekkora szerepe van a folyamatban a népesség korösszetételének, mely rendkívül összetett, hiszen mindig az előző időszak korösszetételéből vezethető le. Ahhoz hogy a folyamatot megfékezzük, a születések számának növelése szükséges, de ennek a népesség növekedése lenne az ára.

A népesség öregedésének számos gazdasági és politikai következménye van: a legfontosabb az idősök ellátásának, gondozásának kérdésköre. Kevésbé szembeutó a folyamat hatása a családok szerkezetére, a generációk arányára, a gazdaság megújuló-képességre és az élet más olyan területeire, amelyekre a korösszetétel változása hatással lehet.

A népesség öregedésének vizsgálatánál fontos megemlíteni a folyamat azon sajátosságát, hogy a rövid és hosszú távú következmények ellentétesek is lehetnek. Az újszülöttek halandóságának csökkenése például a népesség fiatalodásához vezet rövid távon, de ha ezek a generációk felnőttek, akkor a népesség öregedését okozzák. Fordítva viszont ugyanez nem

igaz, tehát az idősök halandóságának csökkenéséből nem következik egyértelműen a népesség fiatalodása.

Az öregedés fogalma az idősök arányának növekedését jelenti a fiatalok arányához képest; a korcsoportok meghatározása és a számítási módszer többféle lehet. Mindenképp szükség van olyan mérőszámra, amely nagyobb adathalmaz áttekintését teszi lehetővé, és összehasonlításra is alkalmas, ez általában egy középérték. A középérték alkalmazásakor mindenképp információvesztés történik. A többféle számítási módszer közötti különbség abban áll, hogy mely információról mond le. A leggyakrabban használt mutatók a népesség öregedésének vizsgálatában az idősök aránya, az életkor mediánja, az egyes korosztályok hányadosai. Emellett egyre gyakoribb az ún. Billeter-arányszám használata is. A különböző mérőszámok kiválasztásánál elsősorban az a döntő érv, hogy mire használjuk fel őket. Ezek röviden a következők.

– Idősök száma (abszolút szám): ekkor az adatok korosztályokba rendezhetők. Egyes kérdéseknél fontos ez a szám, de összehasonlításra nem alkalmas, ahhoz célszerűbb arányszámot használni.

– Medián: ez a szám egy középérték, amely – mint ismeretes – a csoportot mindig két egyenlő nagyságú részre bontja, egy ennél idősebb és egy ennél fiatalabb részre. Ez a szám tehát követi a korcsoport összetételének változását.

– Idősök aránya: (a teljes népességhez viszonyítva): ez a szám az idősök, a 60, illetve 65 év felettiek arányát mutatja a teljes népességhez viszonyítva, tehát az idősök csoportján belüli eltolódásokat „nem érzékeli”.

– Demográfiai időskvóta: megmutatja a már inaktív (60/65 év felettiek) és az aktív korú (20-60/65 éves) népesség arányát, ezért egy ország gazdasági jellemzésénél ez a mutatószám jól felhasználható.

– Billeter-féle arányszám: ez az arányszám megmutatja a generációk egymáshoz való arányát, így a népesség korösszetétel szerint várható fejlődési tendenciáit.

Ha megfigyeljük a fejlett ipari országokban végbement demográfiai változásokat, a következőket

mondhatjuk. Az első fázisban a születési arány magas, míg a halálzási arány jelentősen alacsonyabb. Ez a folyamat hosszú távon a népesség fiatalodásához vezet. A második szakaszban az újszülöttek és csecsemők halálzásának jelentős csökkenése körében a népesség további fiatalodásához vezet, ez a folyamat azonban a harmadik szakaszban megfordul, amikor is a születések száma rendre alulmarad az előző évihez képest, vagyis közép- és hosszú távon ez a népesség önreprodukciójának csökkenését, az átlagéletkor növekedését, az idősök arányának emelkedését jelenti. Ez a folyamat addig tart, amíg a születések számának csökkenése megáll.

Az ENSZ népesség-előrejelzésének 1998-as felülvizsgálata előrevetíti a világ népessége korösszetételének feltűnő mértékű változását, a népesség előrejedését. A korfa 1950-es piramis alakjáról 2050-re átváltozik egy stagnáló ország piramisához hasonló alakzattá. Különösen feltűnő a magasabb korosztályokba tartozók számának relatív növekedése. A népesség öregedése jól látszik a fiatalabb korosztályokba tartozók számának visszaesésében és a halálzási arány csökkenésében is, a várható élettartam majd minden újszülött esetében eléri a 60 évet.

Ha mindezt abszolút számokban vizsgáljuk, ennek az az előnye, hogy mennyiségi betekintést ad a korosztályokba tartozók számába, ami önmagában még nem mond semmit a népesség öregedéséről, de a kérdésfeltevés indikátora lehet, ugyanakkor független a többi korosztályba tartozók számától.

A világ népességének száma 1999-től 2050-ig 6 milliárdról 9 milliárdra nő, a fejlődő országokban élők száma 60 százalékkal, 5-ről 8 milliárdra nő, míg a fejlett ipari országokban a lakosok száma konstans 1 milliárd körüli marad. A 60 és 80 év feletti emberek számának növekedése azonban ennél sokkal dinamikusabb: az 1998-as 580 millióhoz képest 2050-re csaknem 2 milliárd 60 év feletti lakója lesz a Földnek. Ezek az abszolút számok fontos paraméterei az idős emberek gondozásának, az infrastruktúra kiépítésének. Az öregedés elsősorban a fejlődő, illetve a kevésbé fejlett országokban jelent problémát, ahol az alacsony fejlettségi szint mellett gyorsan nő az idős emberek aránya.

Egyedi esetnek mondható Kína, ahol az 1960-as évek „egy gyerek” politikája miatt eleinte gyors öregedés volt megfigyelhető, majd a 2030-as előrejelzés már azt mutatja, hogy a fiatalabb korosztályokba tartozók aránya nem engedi olyan nagy mértékben kiugrani az idősök arányát, mint a többi országban.

Az átlagéletkor változásának figyelemmel kíséréséhez a mediánt elemzi a szerző, ami a népességet két ugyanakkora részre, egy fiatalabbra és egy idősebbre osztja. A fejlődő országokban 1970-ben, a

kevésbé fejlettekben 15 évvel később beállott születésszámú visszaesés értelemszerűen az átlagéletkor növekedéséhez vezetett. A világ minden részére igaz és elmondható, hogy az átlagéletkor a következő 50 évben növekedni fog. A fejlődő országok 2050-re fogják elérni azt az átlagéletkort, ami a fejlett ipari országokra napjainkban jellemző. Így a fejlődő országokban a 60 éven felüli korosztály aránya 2050-re fogja elérni azt a számot, amit a fejlett ipari országok 2000-ben felmutatnak. Ma Dél-Európában a legnagyobb az idősök aránya, 2050-ben pedig az ott élők 40 százaléka 60 év feletti lesz. A „legidősebb országok” Olaszország és Spanyolország lesznek, míg a legfiatalabbak a fejletlen országok, mint például Uganda, Zambia és Jemen, ahol 2050-re ugyanez az arány csak 12 százalék lesz.

Az ún. eltartottsági ráta olyan viszonyszám, amely a gazdaságilag függők – 60 év felettek és 15 év alattiak – arányát mutatja meg a gazdaságilag aktív korú népesség számához képest. Így például ha ez az arány 0,7, amit a fejlett ipari országok 2050-re fognak elérni, akkor ez azt jelenti, hogy 100 aktívra 70 nem aktív ember jut, vagyis ennyit kell eltartania. Jelenleg ez az arány 100 az 50-hez. Ha ennek az arányszámnak az időbeni változását megfigyeljük, akkor azt vehetjük észre, hogy görbéje világszerte egységesen, minden országra jellemzően hullámzik. Ez a születések számának visszaesése miatt alakul így, hiszen ha kevesebb a 15 év alatti lakos, akkor az arány javul, majd később újra romlik, amikor a munkaképes nagyobb generáció nyugdíjas korba kerül.

A Billeter-féle arányszám a generációk egymáshoz viszonyított arányát adja meg. Ha ez az arányszám pozitív előjelű, akkor a gyermek/unokageneráció nagyobb a nagyszülő-generációnál, de lehet negatív előjelű is egy állandó nevezőbeli szülő-generációhoz viszonyítva azt jelenti, hogy a nagyszülő-generáció aránya túllépte az unokagenerációjét. A fejlett ipari országokban ez már az 1970-es évektől kezdődően így van.

Az eddigi vizsgálatok és arányszámok a világ népességének öregedését mutatták ki. Ez a folyamat a fejlett ipari országokban elsőként jelentkezett és 100 évig tartott, míg a fejlődő országokban, élükön Kínával, most zajlik, de előreláthatóan sokkal gyorsabban fog végbemenni. Ennek következményeként a fejlődő országokban az idősöknek egy kevésbé kiépült szociális rendszerben kell majd élniük. A fejlődő országokban tapasztalható halálzási arány visszaesése, ami elsősorban a fiatalabb korosztályt érinti, az öregedés arányát csökkenti, és így módon a születések számának visszaesését ellensúlyozza, hiszen kevesebben halnak meg, így kevesebbnek kell helyettük születni. Ha az öregedés a halálzási arány csökkenéséből ered,

akkor annak egészen más hatásai vannak, például a családok szerkezetére is, mintha a születési gyakoriság csökkenéséből eredne. Előbbi esetben a családok mérete nő, az utóbbiban a család átlagos mérete csökken. Ugyanígy az életkor növekedésével megváltozik a férfiak és a nők aránya is. A 60 év felettiiek esetében, 1998-ban 80 férfi jutott 100 nőre, 80 év felettiéknél

53. A világ népességének öregedése tehát nemcsak makro-, hanem mikroszinten is hatással van az emberek életére. A fejlett országok már kidolgozták stratégiájukat, míg a fejlődőknek minél előbb meg kell tenniük a szükséges lépéseket, hogy meg tudjanak felelni a kihívásoknak.

(Ism.: *Waffenschmidt Jánosné*)

KÜLFÖLDI FOLYÓIRATSZEMLE

Allgemeines Statistisches Archiv

A NÉMET STATISZTIKAI TÁRSASÁG
FOLYÓIRATA

2000. ÉVI 4. SZÁM

Pawlak, M. – Rafajłowicz, E.: Vertikálisan súlyozott regresszió – egy eszköz minőségellenőrző készítéséhez.

Hermann, K. – Lenz, H. J.: Különböző algoritmusok összehasonlítása szóráskomponens-modellek számításához hiányos adatok esetén.

Büchel, F.: Intézményhatás a kiképzésnél, következményei az adatminőségre és az elemzési eredményekre.

Pukelsheim, F.: Mandátumszámok választási arányoknál: a választói szavazatok sikerérték-egyenlősége.

Reich, U. P.: A pénzérték mérése: a tiszta ármozgás statisztikájáról és elméletéről.

2001. ÉVI 1. SZÁM

Saikkonen, P. – Lütkepohl, H.: Egységgyökök tesztelése idősorokban szinteltolás esetén.

Biller, C.: A posteriori módus becslése dinamikus, általánosított lineáris kevert modellekben.

Schrapler, J. P. – Wagner, G. G.: A kérdőbiztos magatartása: leírás és válogatott elemzések a társadalmi-gazdasági panelek „kérdőbiztos paneljeinek” példáján keresztül.

Klein, I.: Ferdeségmérés ordinális skálájú ismérvek esetén rangsorstatistika segítségével.

Schupp, J. – Wagner, G. G.: Szolgáltatások mérése a háztartások kiképzésével.



A FRANCIA STATISZTIKAI
ÉS GAZDASÁGKUTATÓ INTÉZET FOLYÓIRATA

2000. ÉVI 60. SZÁM

Albanesi, S. – Rindi, B.: Az olasz államadósági-kötvény-piac minősége, aszimmetrikus információ és tranzakciós költségek.

Cho, K.H. – El Karoui, N.: Bennfentes kereskedés és nemlineáris egyensúly: egyszerű aukciós eset.

Bisière, C. – Kamionka, T.: A megrendelések időzítése, megrendelési agresszivitás és könyv a párizsi tőzsdén.

Demarchi, M. – Foucault, T.: Kereskedelmi rendszerek Európában: a közelmúlt változásainak áttekintése.

Bauwens, L. – Giot, P.: A logaritmusos ACD-modell.

Benos, A. – Rockinger, M.: Piaci válasz a kereseti bejelentésekre és évközi beszámolókra: az SBF120 társaságok elemzése.

Alphonse, P.: Hatékony árfelderítés a részvényindex pénzállományban és a jövőbeni piacokon.

Ghysels, É. – Gouriéroux, C. – Jasiak, J.: Oksági összefüggés a hozamok és a kereskedés volumene között.

Darolles, S. – Gouriéroux, C. – Le Fol, G.: Napközbeni tranzakciós árdinamika.

Danielsson, J. – De Vries, C. G.: Kockázati érték és extrém hozamok.



A CSEH STATISZTIKAI HIVATAL FOLYÓIRATA

2001. ÉVI 1. SZÁM

Subrtová, A.: A cseh népesedési eszméktől a biológiai és rasszista elméletekig.

Gerylovová, A. – Holčík, J. – Koupilová, I.: A férfiak magas mortalitása Csehország körzeteiben a várható élettartam szempontjából.

Ginter, E.: Az egészségi állapot eltérő alakulása a Szlovák és Cseh Köztársaságban.

Srb, V. – Andrlé, A.: Családi házak – a lakásépítés egy fontos összetevője.

Králová, L.: Házasságon alapuló családok a hagyományos Európában.

Ruzicka, L.: Születésszám-csökkenés Ausztráliában.

Srb, V.: Nemiség (sex) és nem (gender) a demográfiában.



A BIRMINGHAMI EGYETEM FOLYÓIRATA

2000. ÉVI 3. SZÁM

Enev, T. – Koford, K.: A jövedelempolitika hatása az inflációra Bulgáriában és Lengyelországban.

Onour, I. A.: Duális valuta- és devizapiacok egyesítése.

Koop, G. – Osiewalski, J. – Steel, M. F. J.: A kibocsátási szint és a növekedés sztochasztikus határelemzése Lengyelországra és nyugati országokra.

Gaiha, R.: A foglalkoztatottsági garanciarendszer bevezetésének terve az indiai Maharashtra államban.

Jalili, A.R.: Input – output együttműködés aktualizálásánál használt négy, nem adatfelvételeen alapuló technika relatív teljesítményének értékelése.



A FRANCIA GAZDASÁGI
ÉS PÉNZÜGYMINISZTERIUM ÉS A STATISZTIKAI
ÉS GAZDASÁGKUTATÓ INTÉZET FOLYÓIRATA

2000. ÉVI 5. SZÁM

Lagarenne, C. – Legendre, N.: Szegény munkások Franciaországban: egyéni és családi tényezők.

Bontout, O.: A megkeresett jövedelem utáni adóhitelek, adóhitelek az alacsony és közepes jövedelmű háztartásoknak az Egyesült Államokban.

Delarue, V.: A dolgozó családok adóhitele alacsony jövedelmű dolgozó családoknak az Egyesült Királyságban.

Buriccaná, C. – Roth, N.: Az 1912–1914-es generációk életpályájának vége: az intézményi keret hatása.

Magnien, F. – Pougard, J.: Az állandó hasznossági indexek: szintjelző az árnyékvédelem méréséhez.

2000. ÉVI 6. SZÁM

Julien, P.: A városi növekedés mérése és jellemzői.

Bessy – Pietri, P.: A városi növekedés jelenlegi formái.

Brutel, C. – Jegou, M. – Rieu, C.: Területi mobilitás és munkavállalói előléptetés: egy városi térség elemzése.

Rozan, A. – Stenger, A.: A haszonátviteli módszer előnyei és korlátai.

Kremp, E. – Sevestre, P.: Az egy csoporthoz tartozás elősegítő a társasági finanszírozást.



A NEMZETKÖZI STATISZTIKAI INTÉZET
FOLYÓIRATA

2000. ÉVI 3. SZÁM

Rydberg, T.H.: A pénzügyi adatok realizisztikus statisztikai modellezése.

Kokic, P. – Chambers, R. – Beare, S.: Vállalkozási teljesítmény mikroszimulációja.

Victoria – Feser, M.P.: Robusztus módszerek a jövedelemeloszlás, az egyenlőtlenség és a szegénység elemzéséhez.

Schiavo, R.A. – Hand, J.D.: Újabb tíz év a hibaarány kutatásában.

Khuri, A.I.: A varianciabecslés módszerei, múlt és jelen.

Lovison, G.: Általánosított szimmetria-modellek többdimenziós konkordanciatablálknál.

Carlson, B.A.: Nők a statisztikai hivatásban: helyzetjelentés.



AZ AMERIKAI STATISZTIKAI TÁRSASÁG
FOLYÓIRATA

2000. 452. SZÁM

Preisser, J.S. – Galecki, A.T. – Lohman, K.K. – Wagenknecht, L.E.: Dohányzási trendek elemzése hiányos longitudinális bináris válaszoknál.

Hughes, J.P. – Richardson, B.: Egy vertikális HIV-1 transzmisszió megelőzésére szolgáló randomizált kísérlet elemzése.

Auranen, K. – Arjas, E. – Leino, T. – Takala, A.K.: Pneumococcus-hordozás átvitelének a családban: rejtett Markov-folyamatmodell bináris longitudinális adatokra.

Dunson, D.B. – Zhou, H.: Bayesi modell termékenységre és sterilítésre.

Cao, J. – Davis, D. – Wiel, S.W. – Yu, B.: Időváltozó hálózati tomográfia: keretkapcsolati adatok.

Best, N.G. – Ickstadt, K. – Wolpert, R.L.: Térbeli Poisson-regresszió egyenlőtlen bontásban mért egészségi és veszélyeztetettség adatokra.

Sansó, B. – Guenni, L.: Nemstacionárius többhelyszíni modell felhőszakadások elemzésére.

Follmann, D.A.: Leendő kezelést teljesítők közötti kezelési határról: a többszörös rizikófaktor elemzése.

Nandram, B. – Sedransk, J. – Pickle, L.W.: Bayesi elemzés és halálzási ráták leképezése krónikus elzáródásos tüdőbetegségeknél.

Kvam, P.H. – Tiwari, R.C. – Zalkikar, J.N.: Kontaminációs szintek nemparaméteres bayesi becslése reziduális eloszlásokból származó megfigyelések segítségével.

Bayarri, M.J. – Berger, J.O.: P-értékek összetett nullhipotézissel rendelkező (null-modellek) modellek esetén.

Robins, J.M. – Vaart, A. – Ventura, V.: P-értékek aszimptotikus eloszlása összetett null-modellekben.

Cheang, W.K. – Reinsel, G.C.: Autoregresszív becslések torzításának csökkentése idősoros regressziós modellekben korlátozott maximum likelihood segítségével.

Kjell, G.: A szintalapú rétegzett mintavételi terv.

Strawderman, R.L.: Egy növekvő sztochasztikus folyamat átlagának becslése korlátozott leállási idő esetén.

Huang, Y. – Wang, C.Y.: Cox-regresszió pontos segédváltozó nélkül: egy nemparaméteres korrekciós módszer.

Osborne, J.A. – Severini, T.A.: Következtetés egy integrált likelihood függvényen alapuló exponenciális rendstatistikai modellek esetén.

Velasco, C. – Robinson, P.M.: Whittle-féle pseudo maximum likelihood becslés nemstacionárius idősorokra.

Poole, D. – Raftery, A.E.: Következtetés determinisztikus szimulációs modellek esetén: a bayesi egybeolvasztási módszer.

Putt, M.E. – Chinchilli, V.M.: Keresztződő tervek robusztus elemzése többmintás általánosított L-statisztikák segítségével.

Randles, R. H.: Egyszerűbb, affín-invariáns, többváltozós, eloszlásmentes előjel teszt.
Casella, G.: Cimkék a 2000. évre: elmélet és módszerek.
Berger, J. O.: Bayesi elemzés: egy pillantás a mára és a holnap gondolatai.
Brown, L. D.: Tanulmány a statisztikai döntésméletről.
Cappé, O. – Robert, C. P.: Markov-lánc Monte Carlo: 10 év és még fut!
Carlin, B. P. – Louis, T. A.: Empirikus bayesi elemzés: múlt, jelen és jövő.
Christensen, R.: Lineáris és loglineáris modellek.
Efron, B.: A bootstrap és a modern statisztika.
Fan, J.: A nemparaméteres modellezés kilátásai.
Gelfand, A. E.: Gibbs-mintavétel.
George, E. I.: A változó kiválasztásának problémája.
Hettmansperger, T. P. – McKean, J. W. – Sheather, S.J.: Robuszus nemparaméteres módszerek.
Hobert, J. P.: Hierarchikus modellek.
Marden, J. I.: Hipotézisvizsgálat: a p -értékektől a Bayes-faktorokig.
McCulloch, C. E.: Általánosított lineáris modellek.
Meng, X. L.: Hiányzó adatok: Társcsázza M-et, mert???
Portnoy, S. – He, X.: Robuszus utazás az új évezredben.
Reid, N.: Likelihood.
Robins, J. – Wasserman, L.: Feltételképzés, likelihood és koherencia: néhány alapvető koncepció áttekintése.
Solo, V.: Az idősorok vége.
Soofi, E. S.: Alapvető információelméleti megközelítés.
Stefanski, L. A.: Mérésihiba-modellek.
Strawderman, R. L.: Magasabb rendű aszimptotikus közelítés: Laplace, nyeregpont és kapcsolódó módszerek.
Strawderman, W. E.: A minimax elv.

POPULATION

A FRANCIA DEMOGRÁFIAI INTÉZET FOLYÓIRATA

2000. ÉVI 4–5. SZÁM

Munoz-Pérez, F.: Apa nélkül született gyermekek Franciaországban, 1965–1994.
Lemercier, C. – Rosental, P. A.: Falusias vidékek és térbeli megosztásuk.
Sardon, J. P.: Újkeletű változások a fejlett országok demográfiai helyzetében.
Sardon, J. P.: Az 1994-es census a korábbi Jugoszlávia Macedón Köztársaságában.
Dumais, J.: Az új census mintavételi eljárása, becslése és pontossága.
Riandey, B.: Egy éves időszakon keresztül gyűjtött adatok pontossága az új censusban.

REVISTA ROMÂNĂ DE STATISTICĂ

A ROMÁNIAI NEMZETI STATISZTIKAI BIZOTTSÁG FOLYÓIRATA

2000. ÉVI 7. SZÁM

Zamfirescu, A. M.: Július 11. Népszépi Világnap.
Savoitu, G.: Gazdasági és statisztikai kategóriák definiálása árelemzéshez.

Andrei, T.: Strukturális változások elemzése regressziós modellek segítségével.

Jemna, D.: Becslések minőségének javítása az utólagos rétegzési szórás csökkentésével.

Dogaru, V.: A nemzetközi kereskedelem elméleti sémája. A Manoilescu-formula általánosítása.

2000. ÉVI 8. SZÁM

Dinculescu, V.: Az emberi fejlődés mérése és az emberek alapvető jogainak megfigyelése.

Stefanescu, D. E. – Dumitrescu, I.: Nemzetközi esemény – az Európai Statisztikusok Értekezlete.

Botez, O. G. – Galceava, I. B.: Nagy változások a román export- és importstruktúrában a XX. század második felében.

Constantin, D. L.: Idegenforgalom és tartós regionális fejlesztési stratégia. Hatások és költségek.

Pivoda, D.: A fogyasztói árindexek szóródásának statisztikai elemzése.

Parlog, C.: Strukturális változások a munkaerőforrások felhasználásában Romániában a gazdaság átszervezésével összefüggésben.

Aman, O. P.: ARIMA-modellek a román export elemzésében és előrebecslésében.

Troie, L. – Zaharia, O.: Hozzáadottérték-számítás és -elemzés vállalati szinten.

Gheorghiu, P.: A személyzet képzése, szakosítása és javítása a nemzeti statisztikai rendszerben.

Schweizerische Zeitschrift für Volkswirtschaft und Statistik

A SVÁJCI STATISZTIKAI ÉS KÖZGAZDASÁGI TÁRSASÁG FOLYÓIRATA

2000. ÉVI 4. SZÁM

Ungern-Sternberg, T.: Vezetési kompenzációs sémák informáltnál vezetével.

Bonato, D. – Schmutzler, A.: Mikor profitálnak a cégek a környezetszabályozásból? Egy egyszerű mikroökonómiai megközelítés a Porter-féle vitához.

Döpke, J. – Pierdzioch, C.: Tőzsdei szóródás, szektorális megrázkódtatások és a német üzleti ciklus.

Bonjour, D.: Túl vannak-e fizetve a köztisztviselők? Az állami-magán kereseti különbség eloszlásának elemzése Svájcban.

Freitag, M. – Vatter, A.: Közvetlen demokrácia, egyetértés és gazdasági teljesítmény: a svájci kantonok összehasonlítása.

Slovenská štatistika a demografia

A SZLOVÁK STATISZTIKAI HIVATAL FOLYÓIRATA

2000. ÉVI 4. SZÁM

Tirpák, M. – Katerinková, M. – Hecko, I.: Szlovákia népessége 1999-ben a statisztikai adatok tükrében.

Lieskovsky, I. – Papaj, K.: A vállalkozói szféra alakulása Szlovákiában és a közép- és kelet-európai országokban 1995 és 1998 között.

Potocky, R.: Sztochasztikus módszerek a biztosításban.

STATISTICA

Rivista trimestrale fondata da Paolo Fortunati

A BOLOGNAI, PÁDOVAI ÉS PALERMÓI
EGYETEMEK FOLYÓIRATA

2000. ÉVI 2. SZÁM

Maltagliati, M. – Michelini, C.: Ekvivalencia-skálák és fogyasztási egyenlőtlenség: háztartási fogyasztási minták vizsgálata Olaszországban.

Colombi, R. – Forcina, A.: Diszkrét adatok modellezése egyenlőség és egyenlőtlenség kényszerfeltétele mellett.

Arigoni, A.O. – Governatori, L. – Rossi, A.: Egy mesterségesintelligencia-módszer statisztikai adatokból származó információk kinyerésével.

Brizzi, M.: Ferdeség és csúcosság kimutatása robusztus módszerrel: egy új javaslat.

Moustaki, I. – O'Muircheartaigh, C.: Egy egydimenziós latens változómodell nemválaszoláshoz származó, attitűdre vonatkozó következtetéshez nominális adatok esetén.

Nogales, A. G. – Oyala, J. A. – Pérez, P.: Invariancia, majdnem-invariancia és elégségség.

Gupta, A. K. – Kabe, D. G.: Kanonikus korrelációs együttérthetők eloszlása elliptikus hibamodell esetén.

Trivisano, C.: A dolgozó kohorsz mortalitásvizsgálata csoportosított adatokra épített modellel.

Bhargava, M. – Singh, R.: Egy módosított randomizációs eszköz a Warner-modellhez.

Siddiqui, S. R.: Egy rugalmas paraméteres megközelítés a bináris kimenetű modellhez.

Sahoo, J. – Jena, S. J.: A Sen–Midzuno technika hatékonyságáról egy modellfeltételek szerint transzformált változókkal.

STATISTICAL Journal

of the
United Nations Economic Commission for Europe

AZ EGYESÜLT NEMZETEK EURÓPAI GAZDASÁGI
BIZOTTSÁGÁNAK FOLYÓIRATA

2000. ÉVI 2. SZÁM

Fong, P.: Internet használata az Ön térinformatikai rendszerének (GIS) maximalizálásához.

Tammilehto – Luode, M. – Backer, L. – Rogstad, L.: Körzeti adatok és területbehárolás.

Calder, A.: Térbeli elemzés ... Minek vele bajlódni?

Hall, S. – Thurstein – Goodwin, M.: Térinformáció (politikavezérlő – politika által vezérelt), amely statisztikákat szolgáltat az Egyesült Királyság városközpontjai számára.

Rase, D.: A Bizottság kezdeményezései egy európai térinformatikai infrastruktúra létrehozásához.

Nebert, D.: Téradat-kliringház felépítése adatfeltáráshoz és hozzáféréshez.

Grasland, C. – Mathian, H. – Vincent, J.M.: Különálló társadalmi jelenségek többdimenziós elemzése és általánosítása: statisztikai problémák és politikai következmények.

Meliskova, J.: A földrajz és a statisztika integrálása.

Statistical Papers

NEMZETKÖZI ELMÉLETI ÉS ALKALMAZOTT
STATISZTIKAI FOLYÓIRAT

2001. ÉVI 1. SZÁM

Bruner, E. – Puri, M.L.: Nemparaméteres módszerek a faktoriális tervezésben.

Chong, T.T.L.: Változéspontok helyeinek és számának becslése a mintadarabolás módszerével.

Namba, A.: A 2SHI becslés MSE teljesítménye egy regressziós modellben többváltozós t -eloszlású hibaátlaggal.

Cramer, J. S. – Ridder, G.: Máglyázás a multinomiális logit modellben: szabadságfok. Egy korrekció.

Franco, M. – Ruiz, J.M.: Rekordok feltételes várható értékein alapuló diszkrét eloszlások jellemzéséről.

William, M. L. – Durairajan, T. M.: Függvények becslése a bayesi paradigmában.

Wu, J. W.: Geometriai és exponenciális eloszlások általánosított keverékeinek felső extrémális értékeken alapuló jellemzéseiről.

STATISTICAL SCIENCE

A Journal of the American Statistical Association

AZ EGYESÜLT ÁLLAMOK
MATEMATIKAI STATISZTIKAI INTÉZETÉNEK
FOLYÓIRATA

2000. ÉVI 3. SZÁM

Hastie, T. – Tibshirani, R.: Bayesi hátraillesztés.

Stoffer, D.S. – Tyler, D.E. – Wendt, D.A.: A spektrális burkoló és alkalmazásai.

MacKay, R.J. – Oldford, R.W.: Tudományos módszer, statisztikai módszer és a fénysebesség.

Butler, R. W.: Megbízhatóság visszacsatolásos rendszerekhez és nyeregponthoz közelítése.

Morganstein, D. – Marker, D.: Interjú J. Waksberggel.

statistika

EKONOMICKO - STATISTICKY CASOPIS

A CSEH STATISZTIKAI HIVATAL FOLYÓIRATA

2000. ÉVI 12. SZÁM

Drápal, S.: Statisztikai adatgyűjtési program 2001-re.

Rathouská, B.: Alapvető trendek az ipari egységek szerkezetében.

Starik, D.: Idősorok szezonális kiigazítása a Demetra-program segítségével.

Drápal, S.: A lakosság egészsége és a környezet.

2001. ÉVI 1. SZÁM

Bohatá, M.: Egy utasításos rendszer átalakulása és az információ szerepe: a statisztikai szolgálatra vonatkozó néhány következmény.

Jilek, J.: A rövid távú statisztikák havi indexeinek számításai.

Lejnar, J.: Az energiastatisztika harmonizálása.

Kocianová, S. – Král, K.: Információ a „Csehország energiastatisztikájának harmonizálása, 3. fázis – Árak) c. feltáró projekt megvalósításáról.

Novák, I. – Pecáková, L.: Kategorikus változók kapcsolatainak mérése.

2001. ÉVI 2. SZÁM

Soucek, M.: A cseh gazdaság külső környezete – világ-gazdaság 2000-ben.

Jilek, J. – Vojta, M.: A cseh ipar ipari struktúrája 1991-ben és 1998-ban: összehasonlítás.

Tomek, A.: A statisztikai jelentések magyarázó erejéről az építőiparban.

Kuklík, P.: Választások a regionális tanácsokba és a szenátusba 2000. novemberben.

Melléklet: A Cseh Statisztikai Társaság.

Statistische Nachrichten

AZ OSZTRÁK KÖZPONTI STATISZTIKAI HIVATAL
FOLYÓIRATA

2000. ÉVI 12. SZÁM

Népesség-előrejelzés Ausztriára és a tartományaira, 2000–2050.

Demográfiai struktúra és trendek Ausztriában 1999-ben.

Egyeztetett minimálbér-index, 2000. október.

Fogyasztói árindex, 2000. október.

Lakásbérletek, 1998–2000.

Termények és szántóföldek, 2000.

A tartományi kormányzatok energiafogyasztása, 1999.

Idegenforgalom, 2000. szeptember.

Az osztrák nemzeti számlák felülvizsgálata az ESA 95 szerint – Fő eredmények 1998-tól 1999-ig.

Állami deficit – Számítási módszerek és jelenlegi eredmények.

2001. ÉVI 1. SZÁM

Út egy magántársasághoz.

Háztartás- és család-előrejelzés Ausztriára, az 1991–2000. évekre.

Alulfoglalkoztatottság és passzív álláskereső (Munkaerő-felvétel, 1999. március).

Egyeztetett minimálbér-index, 2000. november.

Általános jövedelmi beszámoló, 2000; jogi és módszertani háttér.

Fogyasztói árindex, 2000. november.

A termelő felhasználás felmérése, 1999.

Idegenforgalom, 2000. október.

Idegenforgalom a nyári időszakban, 2000.

Magán- és üzleti utazások 2000. 3. negyedévében.

Keresetiadó-statisztikák, 1999.

2001. ÉVI 2. SZÁM

A 2001-es kombinált census: jogi megalapozás – jelenlegi helyzet.

Kutatás és fejlesztés (R & D) az osztrák vállalatoknál.

Általános jövedelmi beszámoló, 2000; alkalmazottak és nyugdíjasok jövedelme.

Egyeztetett minimálbér-index, 2000. december.

Egyeztetett minimálbér-index 2000-ben.

Fogyasztói árindex, 2000. december.

A települések energiafogyasztása 1999-ben.

Idegenforgalom, 2000. november.

Külkereskedelem 2000. januártól szeptemberig; előzetes eredmények.



AZ OROSZ ÁLLAMI STATISZTIKAI
BIZOTTSÁG FOLYÓIRATA

2000. ÉVI 12. SZÁM

A GDP nemzetközi összehasonlítása Európában.

Grigor'ev, V. V. – Kuznecova, E. V.: A statisztikai adatok védelme az EU országaiban.

Ostapkovitch, G. V.: A gazdasági ciklus mutatóiról.

Burdenkova, E. S.: Statisztika Japánban.

Kbiladze, D.: A vállalatstatisztika reformjának megvalósításáról, megbízható SNA-adatokkal kapcsolatban Grúziában.

Begalov, B. A. – Zaitov, N. – Serikov, M. – Khodiev, B. Yu. – Isakhov, B. A.: Statisztikai osztályozások általános szoftvere.

Az Orosz Föderáció fő társadalmi–gazdasági mutatói 1996 és 2000 között.

Egy teljes körű kisvállalkozási teljesítmény-felvétel eredményeiről 1999 folyamán az Orosz Föderáció 8 régiójában.

Mezőgazdasági háztartások tevékenysége az Orosz Föderációban.

Glisin, F. F. – Kitrar, L. A.: Üzleti tevékenység az Orosz Kereskedelmi Bankoknál: szint és tendenciák.

Frenkel, A. A.: Az orosz gazdaság 2000–2001-ben.

Cejtler, N. V. – Dzhumova, A. V.: A vállalatok külgazdasági tevékenysége a magadani régióban.

Kudrjavceva, R. M. – Matanceva, E. P.: A fafeldolgozó-ipari komplexum jelenlegi helyzete a kirovski régióban.

Wirtschaft und Statistik

A NÉMET SZÖVETSÉGI STATISZTIKAI HIVATAL
FOLYÓIRATA

2000. ÉVI 10. SZÁM

A Szövetség Egészségügyi Monitoring Rendszerének új információs rendszere.

A német külkereskedelem 2000 első félévében.

Az állami munkáltatók alkalmazottjai 1999. június 30-án.

Építési célú megtakarítások, 1999.
Államháztartás 2000 első félévében.
A folyamatos háztartási költségvetési felvétel statisztikai újra átgondolt formában.
A Berlin-módszer egy robusztus adatvezérelt változata.

2000. ÉVI 11. SZÁM

Jelentés a környezetgazdasági elszámolásokról.
Trendek a női foglalkoztatottságban az új tartományokban és Kelet-Berlinben, valamint a Német Szövetségi Köztársaság korábbi területén.
Új trendek az eurózóna országaival folytatott német külkereskedelemben.

2000. ÉVI 12. SZÁM

ÁKM-számítások, 1995 és 1997.
A 2000. évi Gerhard Fürst Díj.

A népesség alakulása, 1999.
Közzszolgálati nyugdíjasok 2000. január 1-jén.
Az építőipar alakulása Németországban 1995-től.
Az idegenforgalmikapacitás-felvétel eredményei, 1999-ben.
Kombinált szállítás.
Árak 2000. november.

2001. ÉVI 1. SZÁM

GDP, 2000.
A népesség alakulása 2050-ig.
Szőlészet Németországban, 1999.
Német és spanyol közösségen belüli kereskedelmi adatok tükrözés összehasonlítása.
Rövid távú bértisztikák Németországban és Franciaországban.
A Birodalom és a német szövetségi állam 100 éves statisztikái.