

MAGYAR SZAKIRODALOM

HUNYADI LÁSZLÓ – MUNDRUCZÓ GYÖRGY –
VITA LÁSZLÓ

STATISZTIKA

Aula Kiadó. Budapest. 1996. 883 old.

Hosszú évtizedeken keresztül a közgazdasági felsőoktatásban részt vevők *Köves Pál* és *Pármiczky Gábor* jegyzeteiből és tankönyveiből tanulták a statisztika módszertanát. Az utóbbi néhány évben azonban örvendetes módon több hasonló igényű jegyzet, illetve tankönyv jelent meg, jelezvén azt a törekvést, hogy a tananyag tükrözze a statisztikai tudományban, a módszerekben időközben bekövetkezett fejlődést, a korszerű ismereteket. Elég e téren utalni az ismertett könyv szerzőinek 1991–1992-ben megjelent *Statisztika I–II.* jegyzetére, illetve a Janus Pannonius Tudományegyetem gondozásában megjelent *Statisztika c. kétkötetes tankönyvre*.¹ A szerzők említett jegyzetének, illetve új könyvének összevetése arra utal, hogy nem egyszerűen a jegyzet könyv alakban történő megjelentetéséről van szó. A Budapesti Közgazdaság-tudományi Egyetem által kiadott, *Köves Pál* által lektorált és *Kerékgyártó Györgyné* által szerkesztett könyv tartalmában kissé bővebb (bár van szűkítés is: a jegyzet Függelékében szereplő levezetések, bizonyítások ugyanis a könyvből hiányoznak), továbbá szerkezetében átdolgozott, a korábbiaknál egységesebb.

Összefoglalóan osztom a szerzők azon reményét, hogy nemcsak oktatási célra használható kitűnően, hanem a statisztikai módszertant alkalmazók is haszonnal forgathatják, akár a különböző módszerek elvi hátterének pontosabb megismerése céljából, akár alkalmazásuk elsajátítására.

¹ *Hajdu Ottó – Herman Sándor – Pintér József – Rappai Gábor – Rédey Katalin: Statisztika. I–II. köt. Janus Pannonius Tudományegyetem. Pécs. 1994. 319 és 245 old. A könyv ismertetése a Statisztikai Szemle. 1995. évi 10. számában (846–850. old.) jelent meg dr. Hunyadi László kandidátus tollából.*

A tekintélyes terjedelmű könyv tíz fejezetre tagozódik, Függelékében pedig a számítógépes outputokhoz angol–magyar szószeret, a valamennyi fejezet végén található gyakorló feladatok közül a fontosabbakhoz megoldást, valamint egyes módszerek alkalmazásához táblázatokat tartalmaz. A könyvet irodalomjegyzék, az angol megfelelőket is tartalmazó tárgymutató, valamint angol nyelvű tartalomjegyzék egészíti ki.

A szerzők hosszú oktatási tapasztalattal rendelkeznek, ez tükröződik mind a könyv felépítésén, szerkezetében, mind abban, hogy az egyes módszerek használatát megvilágító példák, valamint a gyakorló feladatok többsége is érdekes, tanulságos, egyben reális is (kivéve természetesen a mintaelemszámokat, de ez érthető, indokolt). Az esetleges kritikai megjegyzések ugyanakkor kevésbé az oktatási szempontokra koncentrálnak, inkább a felhasználó, a statisztikai módszereket alkalmazó szemszögéből igyekeznek megítélni a könyvet, illetve az egyes fejezetek értékét. Meggyőződésem ugyanis, hogy a könyv megírásába fektetett munka nemcsak a statisztika oktatása terén fog kamatozni, hanem az értő, megalapozott alkalmazások elterjedését is elő fogja segíteni.

A könyv első fejezete öt pontban a statisztikával kapcsolatos alapfogalmakat, az adatszerzési és -hasznosítási módokat, illetve az egyszerű műveleteket, elemzéseket tárgyalja. Talán nem alapvető jelentőségű, mégis elég zavaró a sokaság adott definíciója s főleg a példaként bemutatott több sokaság. Egyrészt már itt célszerű lett volna rámutatni, hogy a fizikai sokaság nem feltétlenül azonos a statisztikai sokasággal, ha egy fizikai sokaság egyedeit nem egy, hanem több számértékkel jellemezzük, ezen különböző számértékek halmazai különböző statisztikai sokaságot jelentenek. Ha például a magyar háztartások mindegyikéhez hozzátartozik egyrészt taglétszámuk, másrészt mondjuk éves hűsfogyasztásuk, ezek a számhalmazok a statisztikai elemzés, a matematikai modell szempontjából két különböző sokaságot jelentenek. De talán

még zavaróbb, ha aggregátumokat nevezünk sokaságoknak. Az 1.1. táblázatban sokaságra példaként említett aggregátumok (például a magyar lakosság húsfogyasztása 1994-ben) nem sokaságok, csak az aggregátum számításánál figyelembe vett egységek (például az egyes háztartások húsfogyasztása) tekinthető sokaságnak. Az adatgyűjtési módszerek között említést érdemelt volna az az egyre szélesebb körben alkalmazott módszer, amelyben az információ kérdőív, szavazólap stb. közbeiktatása nélkül közvetlenül kerül gépi adathordozóra, esetleg rögtön összesítésre. A nem mintavételi hibák között nem szerepel az egyik legfontosabb, a nemválaszolás, továbbá a mintavételi keret hiányosságaiából eredő hibákról sem történik említés. Itt érdemes megjegyezni, hogy nagyon hasznosnak tartom, hogy minden fejezet a tartalomról tömören tájékoztató „felvezető” szöveggel kezdődik, a fejezetek végén pedig van egy rövid összefoglaló is.

A második fejezet az információsűrítés főbb eszközeit mutatja be három pontban. Bár a pontok száma nem utal erre, ez a legbővebb fejezet (125 oldal), amely a sokaságok egy, illetve több ismérv szerinti együttes vizsgálatának standard eszköztárát vonultatja fel. A bő terjedelemben azonban a nem elég visszafogott fogalmazás, a meglehetősen sok ismétlés is hozzájárul. Mivel e téren a gyakorlatban elég sok a téves elképzelés, fontosnak tartom, hogy a könyv hangsúlyozza, hogy egy hisztogram megrajzolásokor nem egyenlő osztályközök esetén nem a téglalapok magasságának, hanem területének kell arányosnak lennie az adott osztályköz gyakoriságával. Az értékösszegekről tárgyalásánál célszerű lett volna megemlíteni, hogy nemcsak az osztályok számától, hanem a vizsgált eloszlás ferdeségétől is erősen függhet az, hogy az osztályközepek és gyakoriságok szorzata mennyire jól közelíti az értékösszegek eloszlását. A koncentráció elemzésével foglalkozó, egyébként nivósan megírt rész végén az R_2 redundanciával kapcsolatban hasznos lett volna rámutatni, hogy az nem más, mint a széles körben használt Theil-féle egyenlőtlenségi mutató (legalább is annak egyik alakja).

A harmadik fejezet – szintén három pontot tartalmazva – részben az elemzéseknél gyakran alkalmazott standardizálás módszerével, részben az indexszámítás kérdéseivel foglalkozik. A standardizálással kapcsolatban hiányolok egy gazdasági jellegű példát, például a férfi és a női átlagbér különbségének felbontását az állománycsoportonként fennálló bérkülönbségek és az állománycsoportok nemenkénti súlya alapján. Tekintettel az indexek fontosságára a hivatalos statisztikában (természetesen másutt is), indokoltnak érzem a kétféle súlyozá-

sú ár-, illetve volumenindexek, ezek előnyeinek, hátrányainak elég részletes tárgyalását, de talán nem ártott volna rámutatni, hogy ha a gyors adatközlésnek van prioritása – mint például a fogyasztói árindexek esetén –, akkor gyakorlatilag csak Laspeyres-típusú index számítása és közlése jöhet szóba. Igen érdekes és sok olvasó számára nyilván újdonság a bilaterális és multilaterális nemzetközi összehasonlításokat szolgáló területi indexekre példaként tárgyalt ECP-project (European Comparison Project), kár, hogy a 3.17 táblában szereplő konkrét adatok kapcsán óvakodik a szerző rámutatni a jelenlegi hivatalos valutaátváltási arányok irreálisára.

A negyedik, a mintavétellel hét pontban foglalkozó fejezettel kapcsolatban mindenképp azt szeretném megjegyezni, hogy számomra logikusabbnak tűnik a mintavétel elméleti és gyakorlati kérdéseit a becslések után tárgyalni, lehetővé téve így a különböző mintavételi eljárásokhoz kapcsolódó becslési módszerek és azok hatásosságának összehasonlítását. Így, a becslésekkel foglalkozó fejezet elé téve, a mintavételi módszerekről inkább csak rövid leíró jellegű ismertetést kap az olvasó, holott manapság már a statisztikai adatok nagy része mintavételből származik, s ezért az adatok megbízhatóságának, statisztikai elemzhetőségének kulcskérdése a megfelelő mintavételi eljárás. Ráadásul még e viszonylag rövid fejezeten belül is elég egyenetlen az anyag. Például a gyakorlatban igen sokszor előforduló többlépcsős mintavételről csak nagyon röviden esik szó, az egyre nagyobb teret kapó nagysággal arányos valószínűségű mintavételről pedig még csak említés sem történik. Véleményem szerint a szórás becslésére szolgáló újabb módszereket (jackknife, bootstrap) nem itt, hanem inkább a becslésekkel foglalkozó fejezetben kellene ismertetni, hiszen ezek nem igazán mintavételi módszerek, s egyébként is ott külön alpont foglalkozik velük. A valószínűségi minta kritériuma szigorúbb, mint amit a könyv megad, bár kétségtelen, hogy legfontosabb eleme az előre rögzített kiválasztási valószínűség a sokaság minden egyedére. A kiválasztási aránynak sokan túlzott jelentőséget tulajdonítanak, ezért nem ártott volna e fejezetben megjegyezni, hogy egyrészt a mintából levonható következtetések megbízhatóságát sokkal inkább a minta elemszáma befolyásolja, mint a kiválasztási arány, másrészt, ha egy rétegzett minta nem arányos eloszlású, nincs is sok értelme a minta egészére vonatkozó kiválasztási aránynak.

A könyv ötödik fejezete nyolc pontban a statisztikai becslések elméleti hátterét és legfőbb módszereit ismerteti. Külön érnyee e fejezetnek, hogy a pontbecslések módszerei mellett részletesen foglal-

kozok az intervallumbecslés elvi alapjaival és különböző eseteivel. Az adott pontosság eléréséhez szükséges mintanagysággal kapcsolatban talán nem ártott volna megemlíteni, hogy a gyakorlatban az alapsokasági szórás ritkán ismert, a relatív szórásra viszont sok esetben reális feltevésekkel lehet élni, s ilyen esetekben hogyan lehet a mintanagyságot meghatározni. E fejezettel kapcsolatban is érzékelhető bizonyos kiegyensúlyozatlanság a tárgyalásban, időnként elég hosszan tárgyal viszonylag triviális dolgokat, fontos részek viszont kimaradtak. Így például véleményem szerint túl van hangsúlyozva a csoportos mintavétel azon speciális esete, amikor a csoportok elemszáma azonos – ez a gyakorlatban csak minőségellenőrzésnél fordul elő –, holott ez, mint a könyv is utal rá, könnyen visszavezethető az egyszerű véletlen mintavételre. Egyáltalán nincs szó viszont sem ebben a fejezetben, sem a későbbiekben például a regressziószámítás kapcsán arról a gyakorlatban szinte mindennapos esetről, amikor a különböző megfigyelésekhez eltérő súlyszám tartozik. Ezt a könyv egyik fő hiányosságának érzem.

A könyv egyik legjobban sikerült részének tartom a hipotézisek vizsgálatával öt pontban foglalkozó hatodik fejezetet, bár néhány elég közismert nemparaméteres próbáról legalább leíró szinten még említett lehetett volna tenni.

Ugyancsak jól felépített, informatív az idősorok elemzési módszereit öt pontban bemutató hetedik fejezet. A módszerek bemutatása során azonban időnként nem ártana olyan alapkérdésekre is felhívni az olvasó figyelmét, mint például értékadatokból álló idősorok esetén, mikor célszerű nominál-, illetve reáladatokot elemezni. Ez utóbbi különösen akkor indokolt, ha a vizsgált időszak alatt jelentősen változott az infláció mértéke. A 7.4 példában is valószínűleg sokkal jobb illeszkedést lehetett volna elérni a reálértékekkel történő számolás esetén, hiszen az 1990-es évek elején megugrott infláció hatását a nominálértékekre nem lehet igazán jól követni egy modellel. A simító eljárásokkal kapcsolatban ellentmondásosnak érzem az α paraméter meghatározását akár a legkisebb négyzetek, akár az abszolút vagy százalékos abszolút eltérések minimalizálásának módszerével, hiszen ezekből szükségszerűen 1-hez közeli α érték adódik, azaz gyakorlatilag nincs simítás.

A szerzők korábbi jegyzetéhez képest jelentős változás, hogy a könyv egy helyett két fejezetben – a nyolcadikban és a kilencedikben – tárgyalja a korreláció- és regressziószámítást, különválasztva a két-, illetve a többváltozós esetet, az előbbivel hat, az utóbbival tizenegy pont foglalkozik. Nem érzem

indokoltnak ezt a szétválasztást, nézetem szerint szerencsésebb lenne a kétváltozós regressziót mint a többváltozós regresszió speciális esetét tárgyalni. Talán csak az egyszerű transzformációval nem linearizálható kétváltozós összefüggés kezelhető nehezen egy többváltozós hasonló modell speciális eseteként, mert ilyen többváltozós modelleknél meglehetősen bonyolult a paraméterbecslés.

Mivel a regressziószámítás területén elég sok formális, nem megalapozott alkalmazással lehet találkozni, nagyon fontos a regressziós modell megválasztása, illetve a különböző modellek valószínűségelméleti háttérének alapos ismerete.

Ez utóbbi szempontból alapvetően három modell különböztethető meg: a magyarázó változók nem, csak a hibátag valószínűségi változó; egy többváltozós valószínűségeloszlás változói közötti feltételes eloszlást fejez ki a regressziós összefüggés; végül az ún. vegyes modell, ahol a magyarázó változók is valószínűségi változók, de szerepel a modellben hibátag is. Ez utóbbi modell azért fontos, mert ha az adatok mintavételből erednek, a magyarázó változók nyilván véletlen változók lesznek, ezek együttes eloszlásáról azonban nem tételezünk fel semmit. A kétváltozós esetet tárgyaló nyolcadik fejezet még külön foglalkozik az X és Y változók együttes (normális) valószínűségeloszlását feltételező modellel – bár a nemlineáris kapcsolatoknál már nem specifikálja, milyen modellről van szó –, a többváltozós regresszió tárgyalása viszont lényegében a nem véletlen magyarázó változók esetre szorítkozik (lásd a 675. old.). Ekkor azonban nem értelmezhetők a magyarázó változók közötti korrelációk, a korrelációs mátrix. Mivel pedig ezek fontos szerepet kapnak a későbbi tárgyalásban, tulajdonképpen az egész fejezet a vegyes modellen alapul. Kár, hogy ez nincs explicit kimondva. A regresszió specifikálása ugyanis véleményem szerint nemcsak a magyarázó változók és a függvénytípus megválasztását jelenti, hanem a modell valószínűségelméleti háttérének egyértelművé tételét is.

A többváltozós regresszióval foglalkozó fejezet egyébként igen gazdagon és alaposan ismerteti a kérdéskörrel kapcsolatos ismeretanyagot, hozzásegítve az olvasót nemcsak az elméleti alapok elsajátításához, hanem a gyakorlati alkalmazáshoz is. Mivel a standard gépi programcsomagok általában tartalmazzák a stepwise módszer mindkét (magyarázó változókat folyamatosan bekapcsoló, illetve elhagyó) változatát, nem ártott volna kicsit részletesebben foglalkozni mással is, mint a backward eljárással, felhívva a figyelmet arra, hogy a kétféle eljárás nem vezet feltétlenül ugyanarra az eredményre, sőt a

végső modell attól is függ, milyen tesztet alkalmazunk egy-egy változó szignifikanciájának eldöntésére. A mul-tikollinearitás problémájának tárgyalásánál célszerű lett volna megemlíteni, hogy ez akkor jelent elsősorban problémát, ha a regressziós modell célja elsődlegesen a különböző magyarázó változók hatásának elemzése. Ha viszont a regressziót elsődlegesen előrejelzésre (vagy interpolációra) kívánjuk használni, akkor kevésbé zavaró a multikollinearitás, ilyenkor a modell minél pontosabb illeszkedése az adatokhoz a döntő. A főkomponens-regresszióval kapcsolatban pedig megemlítendőnek tartanám a főkomponensek közgazdasági (szociológiai stb.) értelmezésének problémáját.

A könyv utolsó, tizedik fejezete három pontban vázolja a statisztika – elsősorban a hivatalos statisztikai szolgálat – szerepét a társadalomban. Egyetértek azzal, hogy nagyon fontos, hogy a leendő közgazdászoknak, de általában is a könyv olvasóinak reális képük legyen a hivatalos statisztikai szolgálat működéséről, céljáról, módszereiről. Talán jó lett volna utalni egyrészt arra, hogy a Központi Statisztikai Hivatal kiadványainak többsége az 1990-es évek eleje óta már kétnyelvű (magyar–angol), másrészt arra, hogy a magyar hivatalos statisztika már nagymértékben igazodik a nemzetközi szervek (ENSZ, EUROSTAT) által kidolgozott fogalmakhoz, módszerekhez, csoportosításokhoz.

Sajnálatos, hogy néhány ábránál (1.2.c, 2.13, 4.2, 4.10, 4.11, 6.4, 7.21) kisebb pontatlanságok vagy a szövegben hibás hivatkozások maradtak, illetve a szövegszerkesztés sem tekinthető tökéletesnek (elírások, sajtóhibák, rossz elválasztások), de ezeket könnyen javítani lehet egy utánnomás vagy a könyv újabb kiadása során.

Bár ismertetésemben több kritikai észrevételt is megfogalmaztam, semmi esetre sem szeretném, ha az olvasóban ezek s nem a könyvet méltató sorok maradnának meg. Az észrevételek elsősorban azt célozzák, hogy a szerzők fontolják meg ezeket, s ha akarják, vegyék figyelembe az oktatásban, illetve a könyv újabb kiadása, esetleges átdolgozása során. Meggyőződésem ugyanakkor, hogy akinek lehetősége nyílik a szóban forgó statisztikakönyvet átolvasni, belőle tanulni, az ismertett módszereket alkalmazni munkája során, tapasztalni fogja, milyen segítséget jelent e kiváló könyv a statisztika elméletének és módszereinek jobb megértéséhez és elsajátításához, amihez a nagyszámú, jól kiválasztott, gyakran gépi programcsomagok alkalmazását is illusztráló példa, gyakorlat is nagyban hozzájárul.

Éltető Ödön

DOBAY PÉTER:

AZ IRODA.
EMBEREK, GÉPEK, SZERVEZET

Panem Kft. – McGraw– Hill Inc. 1996. 295 p.

A könyv szerzője (a Janus Pannonius Tudományegyetem Közgazdaság-tudományi Karának docense) már évek óta tanítja a közgazdászhallgatóknak, a posztgraduális képzésre járó üzletembereknek, kutatóknak, hogy a vállalat irodája, a hivatal sokszor fontosabb, mint a műhely vagy a gépsorok.

A szerző a bevezetőben arról szól, hogy célja összefoglaló ismereteket adni arról, hogyan és miért kell a szellemi munkások (menedzserek, közigazgatási dolgozók, szabad foglalkozásúak: egyre több ember!) munkahelyét hatékonyan megszervezni s kialakítani.

A könyv fejezetei a következők:

1. Az iroda és az automatizálás (információ, folyamatok, technológia);
2. Az iroda funkcionális folyamatai (szövegek, képek, hangok, számok kezelése)
3. Az irodai szervezet (felépítés, költségek, hatékonyság; majd két különleges iroda tárgyalása: az elektronikus dokumentumtovábbítás, az EDI, valamint a távmunka szervezése);
4. Irodai rendszerek szervezése (szükségesség, életciklus, modellek);
5. Emberek és technika (a környezet, a menedzser szerepe);
6. Az iroda jövője (automatizálás, elektronizálás, emberi problémák).

Az irodák a szellemi dolgozók tipikus munkahelyei. Az államháztartási reform, az új vállalkozások, a korszerű közigazgatás, az önkormányzatok olcsó és hatékony működésének igénye, az információs technika szükségessé teszi azt, hogy az irodák hatékonyan működjenek, az információ feldolgozása és továbbítása, az ügyfelek kiszolgálása szakszerű, gyors és megbízható legyen.

A könyv témája pontosan ez: milyen legyen a hatékony iroda, hogyan lehet az irodát automatizálni, melyek az iroda főbb funkcionális folyamatai, mire képes az irodai adatfeldolgozás, az irodai szervezetnek milyen formái vannak, mi az iroda jövője. Ez a témaválasztás a kötet egyik újszerűsége. Jól ábrázolja és bizonyítja, hogy az információfeldolgozási munkahelyek megszervezése, korszerű szemléletű vezetése, technológizálása, igen fontos egy vállalkozás vagy hivatal életében. Mint mindent, ezt is lehet jól vagy rosszul csinálni, de meg is lehet tanulni a hatékony működés technikáit.

Az ügyvitelszervezési és irodatechnikai terület eszközei, módszerei viharos sebességgel avulnak el.